**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application: No. 09/160,657 — Confirmation No. 6611
Filed: September 25, 1998 — Group Art Unit 2822
Inventor(s): Joseph W. Lyding *et al.* — Examiner: VOCKRODT, Jeff B.
For: SEMICONDUCTIVE DEVICES AND METHODS FOR SAME — Attorney Docket: UIL1001-3C

Commissioner for Patents
P.O. Box 1450
Alexandria
VA22313-1450

**CERTIFICATE UNDER RULE 37 CFR 1.10**
I hereby certify, that this document and the documents referred to as enclosed therein are being deposited with the United States Postal Service on the date indicated below, in an envelope as "Express Mail Post Office to Addressee: Mailing Label Number ER 511781828 US addressed to: Assistant Commissioner of Patents, Washington, D.C. 20231.

26 March 2004 — N. Rhys Merrett
Date — N. Rhys Merrett

## DECLARATION OF ROBERT M. WALLACE UNDER 37 CFR 1.132

I, Robert M. Wallace, hereby declare that:

1. I am Robert M. Wallace and my residential address is 2717 Woods Lane, Garland, Texas 75044.

2. I am being paid a fee for my professional services in connection with this Declaration regardless of the outcome of the case and I have no other interest in the outcome of the case.

3. My educational qualifications include a B.S Degree in Physics and Applied Mathematics, a M.S. Degree in Physics and a Ph. D. degree in Physics, all from the University of Pittsburgh.

4. I am presently employed as a Professor in the Departments of Electrical Engineering and Physics of the University of Texas at Dallas, Richardson, Texas 75080.

5. My previous employment has included positions during the period 1990-1999 at Texas Instruments Incorporated, Dallas, Texas as a Member, Technical Staff and Senior Member,

Technical Staff in the Central Research Laboratories, and Manager, Advanced Technology Branch.

6. I am affiliated with various professional bodies, being a Senior Member of the Institute of Electronic and Electrical Engineers, and a member of the Materials Research Society, American Vacuum Society, and Electrochemical Society.

7. I have authored over 60 publications in peer reviewed journals and proceedings and am an inventor or co-inventor in over 35 issued US patents.

8. I am familiar with semiconductor device processing including investigations pertaining to deuterium sintering of MOS transistors as exemplified by publications I have co-authored, including:

"Electrical and physical characterization of deuterium sinter on submicron devices", *Appl. Phys. Lett.*, Vol. 72, No. 14, 6 April 1998, pp. 1721-1723; copy attached marked **"Exhibit A".**

"Hydrogen/Deuterium Interaction with CMOS Transistor Device Structure: Sinteruing Process Studied by SIMS, Mater. Res. Soc. Symp. 513,325 (1998) - 6 pages; copy attached marked **"Exhibit B"**

"Examination of deuterium transport through device structures", *Appl. Phys. Lett.* Vol. 73, No. 23, 7 December 1998, pp. 3441-3443; copy attached marked **"Exhibit C"**.

"Deuterium transport through device structures", *J. Appl. Phys.*, Vol. 86, No. 4, 15 August 1999, pp. 2237-2244; copy attached marked **"Exhibit D"**.

"Deuterium sintering of silicon-on-insulator structures: D diffusion and replacement reactions at the SiO2/Si interface", *J. Vac. Sci. Technol. B,* Vol. 17, No. 5, Sep/Oct 1999, p[p. 2153-2162; copy attached marked **"Exhibit E"**.

9. I have reviewed the specification of US Patent Application No. 09/160,657, filed September 25, 1998, ("Lyding and Hess specification"), together with the claims contained in the Supplementary Amendment dated July 8, 2003, the Office Action mailed September 29, 2003 ("Office Action"), both relating to that application, and International Patent Application WO 94/19829 ("Lisenker") cited in the Office Action. I have been informed that the effective filing date (priority date) of patent application No. 09/160,657 is January 16, 1996, the filing date of patent application No. 08/586,411, of which application No. 09/160,657 is a continuation.

10. The Lyding and Hess specification discloses deuterium annealing of semiconductor devices to improve their operational characteristics, exemplifying reduction in the depassivation of semiconductor devices due to hot-carrier effects (page 8, lines 17-22). Claims 40, 62, 76, 79, and 80 are directed to a semiconductor device which has been subjected to deuterium passivation "post-fabrication" (claims 40, 79), "after formation of at least said gate contact" (claims 62, 80), or "after formation of said source, drain and gate contacts" (claim 76). I note that the specification at page 11, lines 26-28 refers to post fabrication in terms of "subsequent to fabricating the gate, source and drain contacts"; for convenience, I shall refer to these claim recitations as "post-metal annealing". At pages 18-23 of the Lyding and Hess specification, in conjunction with Figs. 2 and 3, an "Experimental" is described in which in a first run wafer samples containing NMOS transistor structures were subjected to post-metal annealing in a 10%/90% ambient of deuterium/nitrogen and in a second run identical wafer samples were subjected to post metal annealing in a 10%/90% ambient of hydrogen/nitrogen under otherwise identical conditions. The description includes results of tests carried out on the annealed wafers under electrical stress and concludes that "transistors sintered in deuterium typically exhibit lifetimes 10 times longer than those sintered in hydrogen." (Page 22, lines 23-27.)

11. Lisenker states: "Higher quality silicon dioxide layers having reduced numbers of dangling and/or weak Si-H bonds are needed to improve the performance of many semiconductor devices. A concomitant new method of fabricatiing such silicon dioxide layers is likewise needed." (Lisenker, page 4, lines 13-17.) Lisenker addresses these needs by teaching, as the Examiner notes, that switching from hydrogen to deuterium in the annealing ambient for the device will incorporate deuterium in the device. In describing embodiments of his invention, Lisenker states: "The formation of Si-D and Si-OD bonds is accomplished in the present invention by contacting a silicon wafer with deuterium or a deuterium containing compound before, during, and/or after formation of a device oxide layer." (Page 6, lines 10-14.) Further, Lisenker teaches (page 7, lines 30-36) that the Si-D bond has a lower zero-point energy than the Si-H bond. This implies that Si-D bond formation is favored over Si-H formation by 71.5- 72.3= -0.8 kcal/mole or only -0.035 eV, as 1 eV is equivalent to 23.06037 kcal/mol. Thus, the Si-D bond is more energetically favorable by only 0.035 eV. Similarly, Lisenker teaches that the Si-OD bond is more energetically favorable than Si-OH formation

102.2-104.3 = -2.1 kcal/mole or only 0.091 eV. (It is noteworthy that a room temperature ambient corresponds to $kT$=0.025 eV or 0.577 kcal/mole, where $k$ is Boltzmann's constant.)

Thermodynamic arguments, as applied by Lisenker, therefore indicate that the Si-D (or Si-OD) bond is slightly less susceptible to bond scission *upon thermal activation* than the analogous Si-H (Si-OH) bond.

Lisenker does not teach or recognize the significance of the resistance of the Si-D bond to *electron stimulated* bond scission, such as that which may arise from hot carrier scattering ("stress"). The arguments offered by Lisenker on the enhanced stability of, for example, the Si-D bond vs. the Si-H bond were based solely on thermodynamic arguments and are most relevant to *thermally activated* bond scission processes. The associated enhanced stability from thermal activation, called a kinetic isotope effect, would be expected to result in a reduction in the rate of Si-D bond scission relative to that observed for Si-H bond scission. The ratio of the rates is typically expected to be proportional to the ratio of the masses – roughly a factor of 2-3. (Lisenker, page 8, lines 3-16.)

On the basis of Lisenker's theoretical discussion, where no data is presented showing device improvements and the small kinetic isotope effect, one ordinarily skilled in the art at January 16, 1996 would not have been motivated to further explore device annealing of a device in a deuterium ambient.

In my opinion, the theoretical teaching in Lisenker's patent application would not have suggested the claimed features of semiconductor devices subjected to post-metal annealing in a deuterium ambient as recited in each of claims 40, 62, 76, 79, and 80 and those claims would not have been rendered obvious to a person of ordinary skill in the art at January 16, 1996 by Lisenker.

12. The Examiner states in the Office Action "Applicants have not shown that annealing post-metallization in deuterium is critical or has unexpected results relative to pre-metal annealing in deuterium (i.e., Lisenker)." I must respectfully disagree.

The data presented in the Lyding and Hess specification (pages 21-23 "ANNEALING RUNS" together with Figs. 2 and 3) ), indicate at least an order of magnitude improvement in device performance – well beyond that taught or anticipated by Lisenker. The demonstrated device lifetime improvement by a factor of 10 or more by Lyding *et al.* was not anticipated in the device community. The Lyding *et al.* publication "Reduction of hot-electron degradation in

metal-oxide semiconductor transistors by deuterium processing," *Appl. Phys. Lett.*, vol. 68, no. 18, p.2526, 1996 (copy attached and marked **"Exhibit F"),** reporting the annealing processing and test results corresponding to those disclosed in Lyding and Hess specification, generated significant interest by others in the technical community, including motivation of myself and my coworkers to examine the improvements in device reliability reported by Lyding, *et al.*

Our initial work, included examination of the impact of a deuterium sinter at 450 °C/60 min, compared to the traditional forming gas (FG) sinter, also at 450 °C/60 min. Channel hot carrier (CHC) measurements indicated that the $D_2$ sinter for 60 min improves the lifetime of the devices by 10× over the FG sinter. DC current–voltage measurements also showed that samples sintered in $D_2$ ambient for 60 min were less prone to degradation under stress. These experiments were conducted in late 1996 and throughout 1997, to the best of my recollection, and our experiments and the results were reported in the publication "Electrical and physical characterization of deuterium sinter on submicron devices", Appl. Phys. Lett., Vol. 72, No. 14, 6 April 1998, pp. 1721-1723 ("Exhibit A").

13. As also noted by the examiner, Lisenker does not teach "post fabrication annealing" in a deuterium ambient.

There have been studies, performed after the original reports by Lyding and coworkers, which have demonstrated that the *retention* of deuterium, and the corresponding *orders of magnitude improvement* in device reliability, depends critically upon the annealing history of the device.

Examples of such studies includes the works I have co-authored entitled "Electrical and physical characterization of deuterium sinter on submicron devices" (Exhibit A), "Hydrogen/Deuterium interaction with CMOS transistor device structure: sintering process studied by SIMS" (Exhibit B), "Deuterium transport through device structures" (Exhibit C), and.).

Specifically, it was shown in these studies that further thermal treatments of a device *subsequent* to a deuterium anneal results in the exchange of deuterium at the interface and throughout the oxide (in the form of, for example, Si-D bonds at the interface or Si-OD bonds in the oxide) with hydrogen. The source of hydrogen may originate from other regions of the device or in the ambient associated with the thermal treatment. Efficient exchange of

deuterium with hydrogen upon thermal treatments was also demonstrated in another study I co-authored entitled "Deuterium sintering of silicon-on-insulator structures: D diffusion and replacement reactions at the $SiO_2$/Si interface" ("Exhibit D).

Studies of the device performance in view of thermal processing and device structures were also published in 1999 by Clark, Ferrence and coworkers. For example, in the article entitled "Process stability of deuterium-annealed MOSFETs" published in IEEE Electron Device Letters (Volume 20, Number 1, pages 48-50 – copy attached hereto as **"Exhibit G"**) it was reported that deuterium initially incorporated from an anneal ($D_2$:$N_2$ = 10%:90%, 400°C, 60 min) and then subjected to a low temperature silicon-nitride deposition process (480°C, 2 min) actually migrates away from the interface.

Taken together, these studies demonstrate that the *retention* of deuterium at the interface necessarily requires the deuterium annealing of the device to be done at a point in the process which precludes the possibility of subsequent migration of deuterium from the gate insulator/semiconductor interface. As such, the recognition of *post-fabrication anneals* in a deuterium ambient uniquely fulfills this requirement. Lisenker does not recognize the significance of the thermal processing history in view of the exchange reactions between, for example, Si-H and Si-D, which would result in rehydrogenation of the interface. Moreover, deuterium anneal processing as described by Lisenker would actually teach away from the beneficial aspects associated with deuterium incorporation as the deuterium would likely migrate away from the interface due to the necessary, subsequent thermal processing required to fabricate the device in an integrated circuit. Prior to the January 16, 1996 priority date of the Lyding and Hess patent application 09/160,657, the importance of the control of deuterium migration in CMOS devices was not understood by those of ordinary skill in the art. These exchange reactions in CMOS device structures were first reported in the work cited above, e.g. Exhibits A, B and C.

14. Lisenker indicates that the deuterium anneal process "is particularly important in those fabrication steps in which a permanent oxide layer is being formed or treated."

However, reflecting the theoretical aspect of Lisenker, the importance of retaining the deuterium at the interface in view of *necessary* thermal treatments after permanent oxide layer formation or treatment are not discussed by Lisenker. As taught by Lisenker, no improvement would have been obtained by employing Lisenker's teaching because of the requirement of

these subsequent thermal treatments associated with device fabrication. A key requirement of Claim 40 is that the FET is "structurally characterized by the *retention* of deuterium" at the channel/gate insulator "interface resulting from post-fabrication passivation" of this interface by heating in a "deuterium gas-enriched atmosphere." (Emphasis added.) Thus, based on the discussion in Section 11 above, that requirement would not have been satisfied by practising Lisenker's teaching. Similarly, the requirement in claim 81 of "a concentration of deuterium introduced into and remaining within" the gate insulator film, "said concentration of deuterium substantially reducing . . . degradation associated with . . . hot carrier stress" would not have been met.

15. The Examiner asserts in the Office Action "Annealing in hydrogen is not the closest prior art. The closest prior art reference is Lisenker." (Emphasis in the original.) This assertion appears not to take into account that whereas Lisenker is a theoretical, paper proposal, post metal hydrogen annealing had been in widespread use in the semiconductor industry for many years – see S. Wolf, "Silicon Processing for the VLSI Era" (Lattice, Sunset Beach, CA, 1995), Vol. 3, p425, 428 (copy attached as **"Exhibit H"**).

Consequently, from the practical point of view of a person of ordinary skill in the art, I consider post-metal hydrogen annealing to be the closest pertinent prior art and a much more appropriate benchmark than Lisenker for Lyding and Hess to have used in carrying out the comparative tests described in the Lyding and Hess specification.

16. W. F. Clark *et al.* in "Improved Hot-Electron Reliability in High-Performance, Multilevel-Metal CMOS Using Deuterated Barrier-Nitride Processing", *IEEE Electron Device Letters,* Vol. 20, No. 10, October 1999, pp. 501-503, (copy attached hereto as **"Exhibit I"**) includes the following commentary:

"Several years ago, the use of deuterium was proposed to reduce NMOS-transistor susceptibility to hot-carrier degradation by hardening the silicon/silicon-dioxide interface." (Citing Lisenker *et al.* Patent Application WO/94/19829, 1994.)

"Recent studies have shown that the replacement of standard hydrogen-based post-metal anneals with deuterium-ambient anneals have yielded significant improvements in hot-electron

lifetimes" – citing Lyding *et al.* "Reduction of hot-electron degradation in metal oxide semiconductor transistors by deuterium processing," (Citing Lyding *et al. Appl. Phys. Lett.*, vol. 68, no. 18, p. 2526, 1996 –"Exhibit F".)

In addition:

"Attempts to anneal in deuterium prior to first-metal have shown significant improvement; however, with further thermal processing, the effect is inherently unstable. – (Citing W. F. Clarke, *et al. IEEE Electron Device Lett.*, vol. 20, pp. 48-50, Jan. 1999 –"Exhibit G".)

These statements by Clarke *et al.* characterize Lisenker and Lyding and Hess in a manner consistent with the characterizations I have made in this Declaration.

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and I declare under penalty of perjury pursuant to the laws of the United States of America that the foregoing is true and correct, and that this declaration was executed by the undersigned on the date indicated next to his signature.

Dated: March 25, 2009

By: Robert M. Wallace

Robert M. Wallace

# Electrical and physical characterization of deuterium sinter on submicron devices

H. C. Mogul,[a] L. Cong, R. M. Wallace, P. J. Chen, T. A. Rost, and K. Harvey
*Texas Instruments, Inc., Dallas, TX 75243*

EXHIBIT A

(Received 16 December 1997; accepted for publication 6 February 1998)

The impact of a deuterium ($D_2$) sinter under two different annealing conditions, 450 °C/60 min and 450 °C/90 min, was studied and compared to the traditional forming gas (FG) sinter. Channel hot carrier (CHC) measurements indicated that while the $D_2$ sinter for 60 min improves the lifetime of the devices by 10× over the FG sinter, an additional increase in the $D_2$ anneal time actually has a negative impact on lifetime. DC current–voltage measurements also showed that samples sintered in $D_2$ ambient for 60 min were the least prone to degradation under stress. Gated diode results showed no appreciable amount of difference in the initial interface state density among the different samples. Secondary ion mass spectroscopy indicated that neither poly nor salicide appears to be a complete barrier to $D_2$ diffusion. © *1998 American Institute of Physics.*
[S0003-6951(98)01314-X]

Degradation of transistor performance due to the injection of energetic electrons has received considerable attention over the past few years. The degradation is believed to occur due to bond breaking by the hot electrons, releasing H from the $Si/SiO_2$ interface. $H_2$ is introduced intentionally during the normal sequence of processing such as sintering the wafers after metal etch, protective oxide (PO) processes, etc. at a temperature around 400–450 °C. This $H_2$ is believed to passivate the dangling bonds at the $Si/SiO_2$ interface thereby reducing the interface trap density and improving the stability of the device.

Recently, it has been demonstrated that the channel hot carrier (CHC) lifetime of *n*-channel metal-oxide-semiconductor (NMOS) transistors annealed in $D_2$ instead of the traditional $H_2$ or forming gas (FG) ambient increases by almost an order of magnitude.[1] It is though that $D_2$ is harder to dislodge under conditions used to desorb hydrogen[1] due to a sizable isotope effect. CHC lifetime results reported in this work use silicon-nitride sidewall and are salicided in contrast to previously reported results.[1,2] Additionally, we also investigate the effect a longer $D_2$ sinter time on the device properties.

The devices used in this study were NMOS transistors with a nominal 55 Å gate oxide for a 2.5 V technology. Wafers were removed after metal-1 etch and split into three groups (i) FG (10% $H_2$/90% $N_2$) sinter at 450 °C/60 min, (ii) $D_2$ (10%$D_2$/90% $N_2$) sinter at 450 °C/60 min, and (iii) $D_2$ (10%$D_2$/90% $N_2$) sinter at 450 °C/90 min. Following the sinter steps, CHC measurements were performed to study the impact of the sinter conditions on lifetime of devices. Gated diode measurements were performed to obtain the initial (time-zero) interface state density while dc current–voltage (DCIV) tests were done to study the damage at the $Si/SiO_2$ interface after the stress. Physical characterization was subsequently performed by the dynamic secondary ion mass spectroscopy (SIMS) technique.

[a]Electronic mail: h-mogul@ti.com

CHC measurements were performed by stressing the devices under different conditions, namely, 3.75 V/0.1 h, 3.45 V/1 h, and 3.05 V/10 h at peak substrate current ($I_{sub}$) conditions. Degradation to the interface, as measured by shifts in various transistor parameters such as transconductance ($g_m$), drive current ($I_{ds}$), etc., was monitored over the stress time period. Time to fail criteria was based upon time required for each of the parameters monitored to degrade by 10% of the initial values.

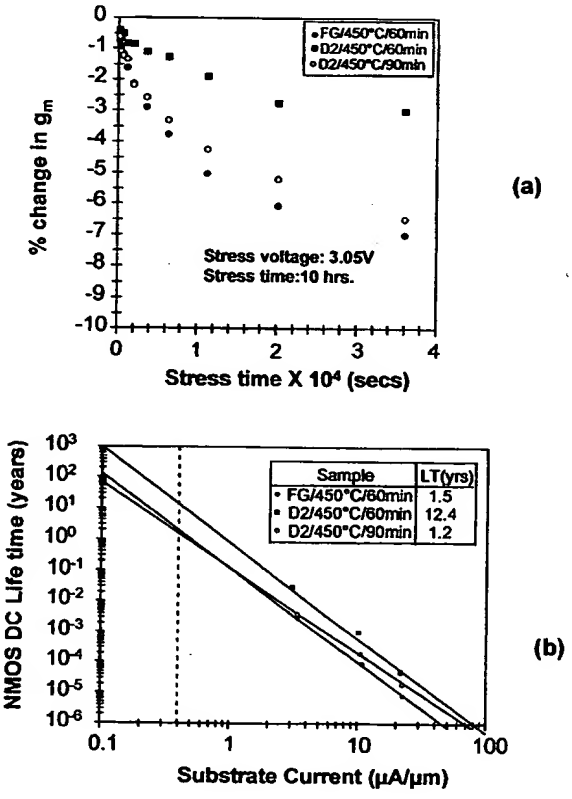


FIG. 1. (a) Time-dependent degradation of $g_m$ for NMOS devices annealed in FG and $D_2$ for two different sinter times. (b) CHC lifetime based upon 10% change in $g_m$ for NMOS devices annealed in FG and $D_2$ for two different sinter times.

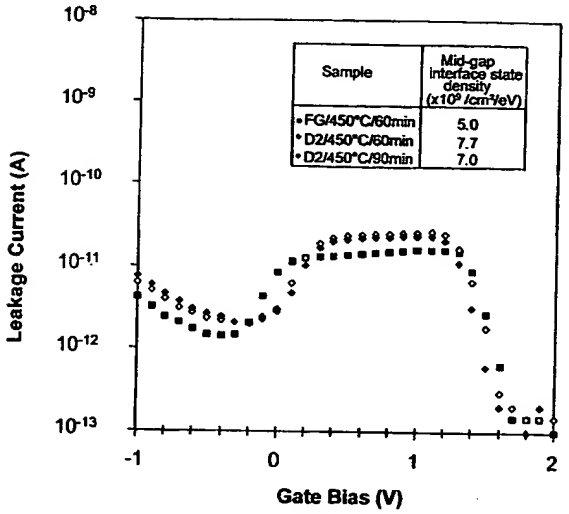


FIG. 2. $n^+/p$ gated diode junction leakage at a reverse bias of 1 V as a function of gate voltage for samples annealed in FG and $D_2$ for two different sinter times.

Figure 1(a) is a representative figure that shows the $g_m$ degradation for devices stressed at 3.05 V for 10 h. The figure clearly indicates that the sample with $D_2$ sinter for 60 min degraded less than the FG sample which was annealed under identical conditions. However, when the $D_2$ sinter time was increased to 90 min we notice that the degradation rate did not improve when compared to the $D_2$ sample sintered for 60 min. These initial results indicate the need for optimizing the sinter conditions to obtain the maximum benefits from $D_2$ anneal. Figure 1(b) shows the lifetime of the NMOS transistors sintered in $D_2$ for 60 min have a 10× improvement over those sintered in FG or $D_2$ for 90 min. Other transistor parameters such as $I_{ds}$ show a similar trend. Each data point in Fig. 1(b) represents an average of several measurements.

Comparison of $n^+/p$ gated diode leakage at a fixed reverse bias of 1 V, as function of gate voltage for different samples is indicated in Fig. 2. Mid-gap interface trap density ($N_{it}$) was calculated[3] and is shown in the inset of Fig. 2. The calculated trap density indicates no appreciable difference among the different samples. This result indicates that $H_2$ or

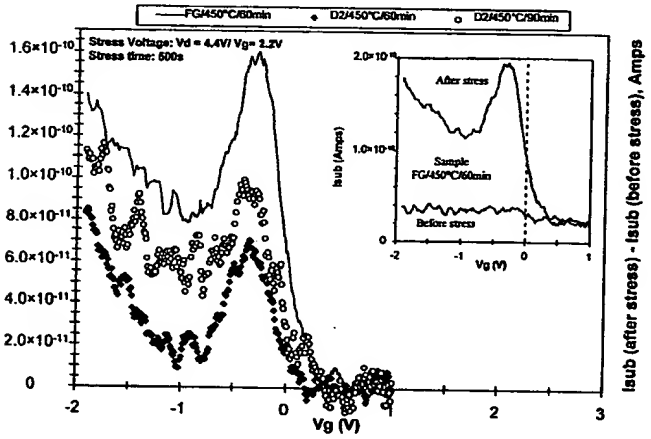


FIG. 3. Comparison of the difference between substrate currents (before and after stress) from DCIV measurements for NMOS devices stressed at $V_{ds}/V_{gs}$=4.4 V/2/2 V for 500 s for samples annealed in FG and $D_2$ for two

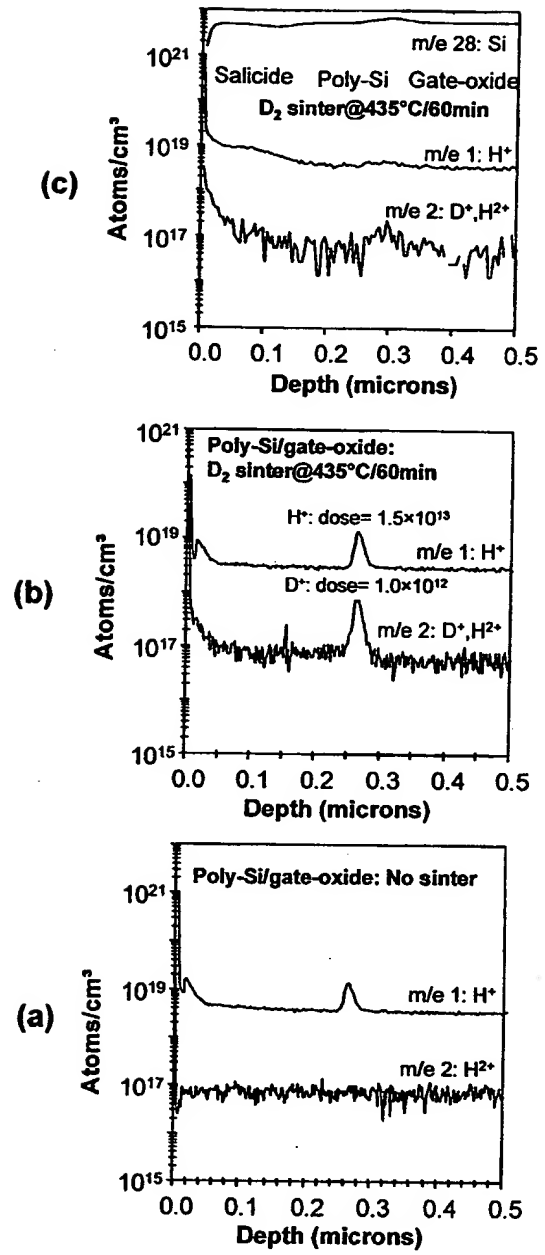


FIG. 4. (a) Dynamic SIMS depth profiles for poly-Si/gate-oxide structure before $D_2$ sinter (b) Dynamic SIMS depth profiles for poly-Si gate-oxide structure after 435 °C/60 min $D_2$ sinter. (c) Dynamic SIMS depth profiles of $TiSi_2$/poly-Si/gate-oxide capped oxide structure after 435 °C/60 min $D_2$ sinter.

$D_2$ ambients are indistinguishable for $N_{it}$ passivation at time zero. The real impact of $D_2$ annealing occurs after the devices are subjected to an electrical stress and this is clearly indicated in Fig. 3 (inset) with the DCIV measurement[4,5] performed on NMOS transistors. In the DCIV test $I_{sub}$ is monitored as a function of gate voltage and as such an increase in $I_{sub}$ is an indication of hot carrier induced stress related defects that lie at the interface. Figure 3 clearly indicates that the difference in $I_{sub}$ before and after CHC stress is higher for the FG sample than that for the $D_2$ sample. This result complements the previous CHC result in indicating that the $D_2$ sample sintered for 60 min provides better immunity towards interface damage.

Dynamic SIMS analysis was performed to measure the

performed on structures (350 $\mu$m×350 $\mu$m) specifically designed for SIMS with an $O_2^+$ ion beam at 8 keV. Positive secondary ions ($Si^+$, $H^+$, $D^+$) were collected as a function of sputter time and atomic H/D concentration quantification in Si is achieved by using a H-implanted standard.[6] Figures 4(a) and 4(b) show the SIMS profile for samples before and after the $D_2$ sinter (435 °C/60 min) on poly/gate-oxide stack. Figure 4(c) shows the results for a salicide/poly/gate-oxide stack. From Fig. 4(b) it is evident that $D_2$ diffusion can occur through the structure allowing it to reach the $Si/SiO_2$ interface leading to the formation of stable Si-D species. The quantified Si-D areal density was observed to be $10^{12}$ $cm^{-2}$ which is in contrast to the value of $10^{14}$ $cm^{-2}$ published in Ref. 2. Such a high value as reported in Ref. 2 would appear inconsistent with expected interface state densities. Figure 4(c) shows a similar Si-D enrichment on the salicide/poly/gate-oxide stack despite the degraded depth resolution from the ion-sputtered induced roughness on salicide. This figure also illustrates an important result, namely, that metallized NMOS gate ($TiSi_2$/poly) may not be a hindrance to $D_2$ diffusion. We have also examined poly/gate-oxide stack structures after $D_2$ sinter for 90 min and no increase in [D] was observed.

In conclusion we find that $D_2$ sintered samples are less prone to degradation due to electrical stress than FG samples even for the case of salicided samples using a nitride sidewall technology. We also observe that a longer $D_2$ sinter time does not necessarily provide an improvement in immunity from electrical stress. SIMS results indicate that salicide does not appear to hinder $D_2$ diffusion. It is also important to realize that an improvement in the transistor lifetime with the $D_2$ anneal may not necessarily translate to an improvement in the circuit lifetime. This is because the circuit may have design complexities which override single transistor issues.

[1] J. Lyding, K. Hess, and I. Kizilyalli, Appl. Phys. Lett. 68, 2526 (1996).
[2] I. Kizilyalli, J. Lyding, and K. Hess, IEEE Electron Device Lett. 18, 81 (1997).
[3] H. C. Mogul, T. A. Rost, and D. G. Lin, IEEE Trans. Electron Devices 44, 388 (1997).
[4] C. T. Sah, IRE Trans. Electron Devices 9, 92 (1962).
[5] C. T. Sah, A. Neugroschel, K. Han, and J. Kavalieros, IEEE Electron Device Lett. 17, 72 (1996).
[6] H-implant in Si: standard reference material, Charles Evans & Associates, Redwood City, CA.

in *Hydrogen in Metals and Semiconductors, Materials Research Society Symposium Proceedings* **513** (1998) 325

# HYDROGEN/DEUTERIUM INTERACTION WITH CMOS TRANSISTOR DEVICE STRUCTURE: SINTERING PROCES



P.J. CHEN and R. M. WALLACE
Si Technology Research Center
Texas Instruments Inc., 13536 N. Central Expressway, M/S 921, Dallas, TX 75243

## ABSTRACT

Passivation of the $SiO_2$-Si interface by hydrogen/deuterium in MOS transistors serve to ensure their operating reliability against channel hot carriers. Physical characterization of device sintering process in deuterated forming gas ($10\%D_2$:$90\%N_2$) is carried out by dynamic SIMS on planar CMOS gate stack structures, in conjunction with device hot carrier electrical testing. It is found that incorporation of deuterium in the doped poly-Si/$SiO_2$/Si interfacial region readily occurs under typical post-metallization sintering conditions, demonstrating that transport of deuterium through CMOS gate is an effective pathway in an encapsulated device structure with silicon nitride sidewalls. The measured Si-D areal densities in the interfacial region depend on gate poly-Si doping type, but in both cases, appear to be sufficient to achieve complete interface Si dangling bond ($\sim 10^{12}$ $cm^{-2}$) passivation for the $SiO_2$-Si system.

## INTRODUCTION

Hydrogen passivation of the $SiO_2$-Si interface states and its robustness under device operation conditions play a key role in assuring MOS transistor channel hot carrier reliability. Recent discovery of a marked isotope effect by replacing hydrogen with deuterium in the post-metallization sintering process has sparked much renewed interest in this technologically important area and raised some fundamental materials and physics questions [1-4]. These device studies [1,2,4-6] have indirectly confirmed the beneficial effect of deuterium passivation of the $SiO_2$/Si interface. Thus, in conjunction with electrical hot carrier reliability testing, it is of key importance to be able to conduct direct physical measurements on relevant device structures to confirm interfacial Si-D enrichment after deuterium sintering and establish optimum conditions for such a process [5,6]. Published device studies to date which show marked improvement after deuterium sintering have been largely obtained from NMOS devices with oxide sidewall spacers. Diffusion of hydrogen or deuterium in Si oxide is facile under typical sintering conditions, as can be inferred from many early interface passivation studies with thick (≥1000Å) gate $SiO_2$. It has been suggested [4,5] that the nitride sidewall may serve as an effective barrier for H or D diffusion to the $SiO_2$/Si interface (Fig. 1). Alternatively, H or D diffusion through the reoxidized poly-Si gate sidewalls may well serve as effective diffusion pathways for H or D delivery to the $SiO_2$/Si interface, regardless of the presence of the nitride sidewalls. Finally, the diffusion of H or D through the poly-Si gate (perhaps through grain boundaries) also remains a possibility for H and D transport.

The purpose of this work is to study the interaction of deuterium *molecule* with a planar MOS transistor structure at low temperatures (<500°C) using secondary ion mass spectrometry (SIMS), with particular emphasis on the deuterium transport mechanism and formation of Si-D at the gate $SiO_2$/Si interface. In conjunction with device electrical characterization, this study

in *Hydrogen in Metals and Semiconductors, Materials Research Society Symposium Proceedings* **513** (1998) 325

will help to further elucidate potential deuterium transport pathways relevant to nitride sidewall CMOS device structures.

Our experimental results suggest that the interaction of *molecular* hydrogen or deuterium with the poly-Si/$SiO_2$/Si system is rather complex and highly dependent on the physical properties of the polysilicon itself. Under typical low temperature (<500°C) sintering conditions, the transport of deuterium through a CMOS transistor gate stack is at least competitive (if not dominating) with other suggested diffusion pathways, such as through oxide sidewall spacers.

## EXPERIMENT

Planar CMOS test blocks of 500x450μm$^2$ were used for SIMS analysis. These test blocks were fabricated along with transistor test structures using standard CMOS process, with a nominal 55 Å gate $SiO_2$, complementary p+/n+ doped poly-Si gate (in ~$10^{20}$ cm$^{-3}$ range), and silicon nitride sidewall spacers. Wafer samples for SIMS were sintered in deuterated forming gas (10%$D_2$:90%$N_2$) at 435°C for 1 hour after source-drain anneal and after gate metallization ($TiSi_2$) process steps, respectively. Wafers for electrical testing received further processing (up to metal-1) before the final sintering in deuterated forming gas.

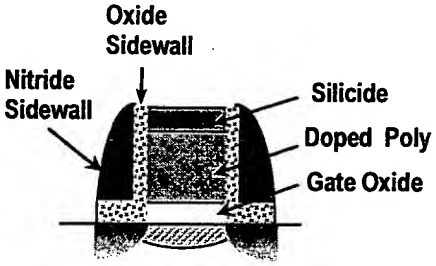


Fig 1. Schematics of a MOS transistor gate structure

Dynamic SIMS depth profile analyses were carried out using both $O_2^+$/positive secondary ion and $Cs^+$/negative secondary ion modes, respectively, on a magnetic sector SIMS instrument. A silicon wafer ion-implanted with $^1H^+$ and $^2H^+$ was used as calibration standard for determining hydrogen/deuterium concentration in Si.

## RESULTS AND DISCUSSION

### SIMS Characterization of Planar CMOS Device Structures

Fig. 2 show SIMS depth profiles of PMOS and NMOS poly-Si gate stack before and after deuterium sinter. It is quite evident that following sintering in $D_2$:$N_2$ forming gas at 435°C, deuterium enrichment is taking place in the poly-Si/$SiO_2$/Si interfacial region. This provides direct evidence that deuterium diffusion/transport through doped poly-Si gate can readily occur under current sintering conditions, allowing deuterium to reach the Si/$SiO_2$ interface region and interact with interfacial reactive sites, such as Si dangling bonds (designated as Si•).

It is noted that deuterium is also incorporated in the PMOS boron-doped p+ poly-Si gate, at a quantified bulk concentration level of ~2-3x$10^{18}$ cm$^{-3}$. This interesting differences between p+ and n+ poly-Si is further highlighted by the negative secondary ion SIMS depth profiles shown in Fig. 3. Owing to the complete suppression of molecular ion ($H_2$) background in the negative secondary ion mode, it is evident that the NMOS phosphrous-doped n+ poly-Si gate is also incorporated with deuterium, at a somewhat lower concentration level of ~$10^{17}$ cm$^{-3}$.

Negative ion SIMS from silicided ($TiSi_2$) poly-Si gate are shown in Fig. 4. The results confirm that the metalized ($TiSi_2$) poly-Si gate does not form a barrier to deuterium diffusion. The observed deuterium distribution in the PMOS and NMOS poly-Si/$SiO_2$ layers below $TiSi_2$ is

very similar to that of doped poly-Si only case, with some additional peak broadening due to ion sputtering induced roughening effects in the $TiSi_2$.

To obtain a meaningful estimate of deuterium areal density accumulated in the poly-$Si/SiO_2/Si$ interfacial region, integration of deuterium concentration over the entire interfacial region need to be performed. We choose the positive secondary ion SIMS data (Fig. 2) for this purpose since the ion yield variations across the interface region simply follow the local oxygen concentration and sensitivity factor correction (for H in Si vs. $SiO_2$) can be performed to limit the uncertainly in concentration to within a factor of 2 (or 100%) [7]. For the NMOS gate, the deduced deuterium areal density is $4x10^{12}$ $cm^{-2}$ and for the PMOS gate, $3x10^{13}$ $cm^{-2}$. Despite the limited accuracy in these numbers, both values appear more than sufficient for what's required to achieve complete interface Si dangling bond passivation (commonly at ~$1x10^{12}$ $cm^{-2}$ levels).

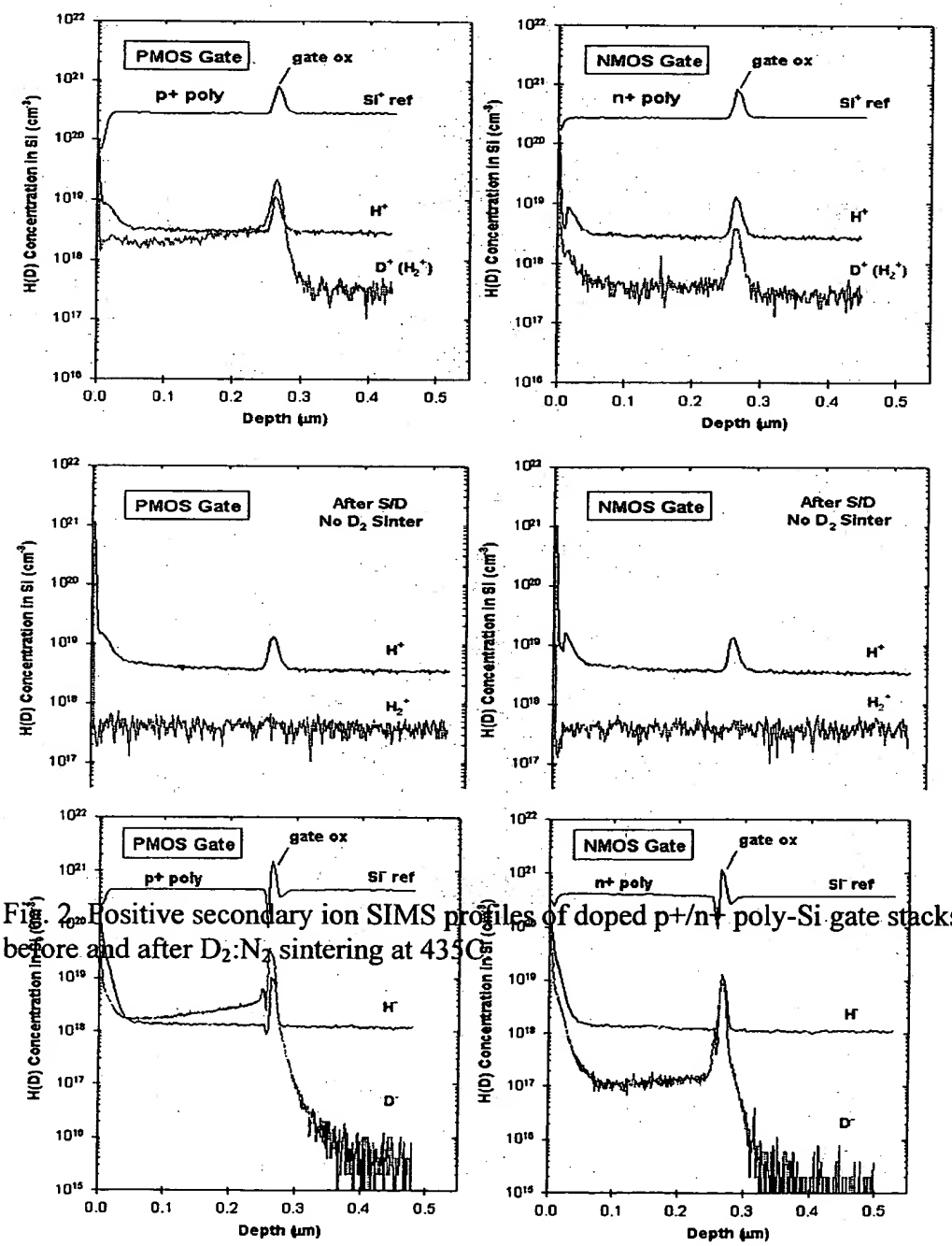


Fig. 2 Positive secondary ion SIMS profiles of doped p+/n+ poly-Si gate stacks before and after $D_2$:$N_2$ sintering at 435C.

Fig. 3 Negative secondary ion SIMS profiles of doped p+/n+ poly-Si gate stacks

in *Hydrogen in Metals and Semiconductors, Materials Research Society Symposium Proceedings* **513** (1998) 325

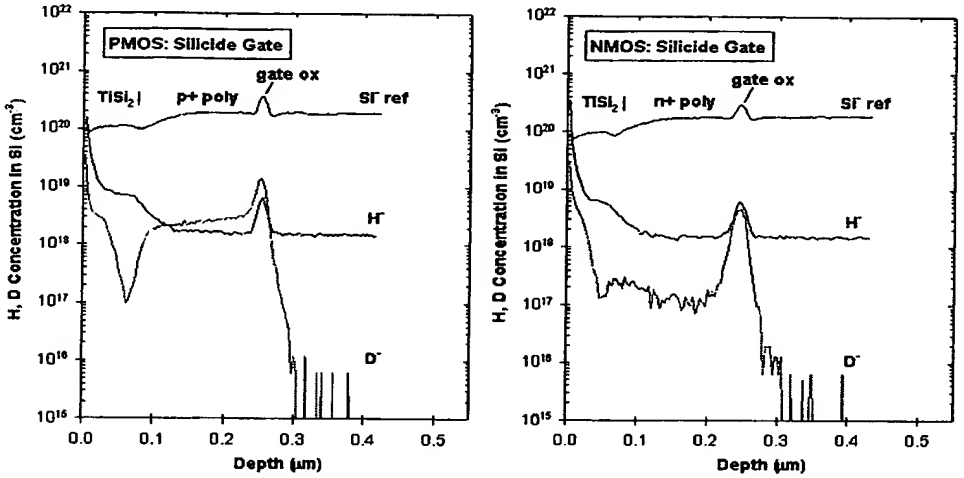


Fig. 4 Negative secondary ion SIMS profiles of CMOS Ti-silicide/poly-Si gate stack after $D_2$:$N_2$ sinter

NMOS Device Reliability Against Channel Hot Electrons

We have reported elsewhere [6] the electrical test results of NMOS transistors fabricated with standard silicon nitride sidewall technology. These NMOS devices received further processing than the SIMS samples up to metal-1. To briefly summarize the electrical results here, a factor of 10x improvement in NMOS lifetime (10% degradation definition) against channel hot electrons was observed for devices sintered in $D_2$:$N_2$ at 450°C compared to $H_2$:$N_2$.

Effect of Extended Deuterium Anneal

According to our dynamic SIMS measurements, wafers subjected to additional $D_2$:$N_2$ sinter at 435C/1hour conditions yielded *the same or less* Si-D population in either NMOS or PMOS. From a process engineering point of view, this suggests that the initial 435C/1 hour $D_2$:$N_2$ sinter is sufficient to reach optimal Si-D population in these particular device structures.

Hot carrier reliability measurements on NMOS devices which have endured a $N_2$:$D_2$ sinter for times > 1hour have not shown the same performance improvement [6]. It suggests that a competitive interface reaction, possibly Si-H replacement due to the substantial H concentrations in the device, might occur during extended anneal.

Deuterium Transport Mechanism and $SiO_2$/Si Interface Passivation

Presence of dopants in the poly-Si undoubtedly plays a major role in promoting deuterium transport through the CMOS gate stack, according to the SIMS results. In crystalline Si, atomic hydrogen(deuterium) is known [8] to form chemical complex with both boron and phosphorus dopants, which results in dopant electrical deactivation at modest temperatures (≤300°C). This effect is much more pronounced for boron dopants compared to phosphorus, where the hydrogen can be in a positively charged state in the complex, Si:B-H. We suggest that similar dopant-mediated hydrogen(deuterium) complex formation also occurs in the CMOS doped poly-

Si gates during the sinter process and is responsible for the deuterium distribution profiles in the poly-Si as measured by SIMS. This process serves to increase hydrogen(deuterium) solubility in Si and therefore enhance its transport to the $SiO_2$/Si interface by normal (interstitial) diffusion channel. Additionally, grain boundary diffusion in poly-Si may also be an effective channel in delivering deuterium to the gate $SiO_2$/Si interface. At present, precisely how *molecular* hydrogen (deuterium) interacts with dopants in the poly-Si is unclear and a specific mechanism can not be deduced from the limited data here. The effect of dopant concentration on hydrogen transport in poly-Si is currently the subject of further investigation.

The diffusion of hydrogen in crystalline silicon is a subject of significant complexity because hydrogen can be present in several forms ($H_2$, $H^0$, $H^{\pm}$) and interact strongly with impurities or defects [8,9]. In our present work, since the observed passivation effects originates from *molecular deuterium* ($D_2$), it's transport through poly-Si to the $SiO_2$-Si interface will *necessarily* involve a dissociation step at one point. Based on published studies [8,9], the interstitial diffusion of *atomic hydrogen* ($H^0$) in mono-crystalline Si is extremely facile in the sintering temperature range (400-600°C). Thus, the rate-determining step to $SiO_2$-Si interface passivation is most likely the dissociation of the $H_2$ molecule. Presence of dopants or Si dangling bonds in poly-Si will profoundly modify the formation mechanism of the hydrogenous species ($H^0$, $H^{\pm}$) from molecular hydrogen (or deuterium) and most likely enhance their transport through the poly-Si.

It is important to recognize this parallel pathway for hydrogen(deuterium) transport through the doped poly-Si gate stack, particularly for a nitride-sidewall CMOS structure. The situation can be quite different for oxide-sidewall CMOS structures, where *molecular* hydrogen ($H_2$) has been shown to effectively transport through the silicon dioxide network at relatively low temperatures (225°C and up) and dominate the interface passivation process [10].

Presence of hydrogen-containing ambient is quite prevalent in CMOS backend process and there is abundance of hydrogen in interlevel dielectrics. From a device process engineering point of view, much parameter space need to be explored in order to maintain and maximize the desired beneficial effects of Si-D passivation of the $SiO_2$/Si interface.

## CONCLUSIONS

Based on the results of SIMS studies, we have demonstrated that:

1. Molecular deuterium interacts strongly with CMOS doped poly-Si gate material in a low temperature sinter (<500C) process.
2. Passivation of the $SiO_2$/Si interface by Si-D can be readily achieved as evidenced from SIMS measurements.
3. Neither doped poly-Si, nor metalized poly-Si gate ($TiSi_2$) is a barrier to deuterium diffusion.

These results further suggest that transport of deuterium through a CMOS gate stack is efficient and competitive with other potential diffusion/transport pathways. They also provide a good physical basis for the observed transistor electrical reliability improvement against channel hot electrons in nitride sidewall NMOS devices.

## ACKNOWLEDGEMENTS

in *Hydrogen in Metals and Semiconductors, Materials Research Society Symposium Proceedings* **513** (1998) 325

We acknowledge the support of DP1 Productization Group, Jiejie Xu , Ken Harvey in supplying the fabricated samples used in this study and valuable discussions with Homi Mogul and Ling Cong about electrical measurement results.

REFERENCES

1. J. Lyding, K. Hess, and I. Kizilyalli, Appl. Phys. Lett. **68,** 2526 (1996).

2. Kizilyalli, J. Lyding and K. Hess, IEEE Electron Devices Lett. **18**, 81 (1997).

3. C. G. Van de Walle and W. B. Jackson, Appl. Phys. Lett **69,** 2441 (1996).

4. K. Hess, I. Kizilyalli, and J. W. Lyding, IEEE Trans. Electron Devices **45,** 406 (1998).

5. J Lee, S. Aur, R. Eklund, K. Hess and J.W. Lyding, J. Vac. Sci. Technol., in press.

6. H. C. Mogul, L. Cong, R. M. Wallace, P. J. Chen, T. A. Rost, K. Harvey, Appl. Phys. Lett. **72,** 1721 (1998).

7. The ion yield increase of $H^+$ in $SiO_2$ vs. Si matrix tend to be offset by the rise in $Si^+$ ion reference signal, based on our measurement of $H^+$ implanted in $SiO_2$ and Si, respectively. The same trend is assumed to hold true for $D^+$.

8. S. J. Pearton, J. W. Corbett and M Stavola, Hydrogen in Crystalline Semiconductors Springer-Verlag (1992), and extensive references therein.

9. J. Chevallier and M. Aucouturier, "Hydrogen in Crystalline Semiconductors", Ann. Rev. Materl. Sci. **18**, 219 (1988).

10. K. L. Brower, "Passivation of Paramagnetic Si-$SiO_2$ Interface States with Molecular Hydrogen", Appl. Phys. Lett. **53**, 508 (1988).

# Examination of deuterium transport through device structures

P. J. Chen and R. M. Wallace[a)]
*Texas Instruments, Inc., Dallas, Texas 75243*



(Received 7 July 1998; accepted for publication 2 October 1998)

We use secondary ion mass spectrometry to characterize the hydrogen/deuterium distribution and concentration on 0.18 $\mu$m "metal" oxide silicon test structures subjected to deuterium anneals. We examine the temperature dependence and the influence of doping on the transport of deuterium to the gate oxide interfaces resulting in interface passivation. We find that *undoped polycrystalline silicon* appears to be an efficient barrier for deuterium transport at typical postmetallization sintering temperatures. © *1998 American Institute of Physics.* [S0003-6951(98)00149-1]

Passivation of $SiO_2$–Si interface states through a "back end" process with hydrogen anneals (sinters) of metal oxide semiconductor field effect transistor (MOSFET) devices is a common and critical step in obtaining reliable, high performance.[1] Additionally, recent reports have shown that *n*-type MOS (NMOS) channel hot carrier (CHC) lifetime can be markedly improved (10–50×) by low temperature (400–450 °C), postmetallization (metal-1) sinter in *molecular deuterium* ($D_2$).[2–5] In these processes, the efficient transport and delivery of such hydrogenous species to the $SiO_2$/Si interface is important to obtain the desired improvement. In this work, we examine the temperature dependence and the influence of polycrystalline silicon (poly-Si) doping on the transport of deuterium to the gate oxide interfaces arising from a $D_2$ sintering. One of the key findings here is that *undoped poly-Si* forms an efficient barrier to deuterium transport at typical sintering temperatures (~400–480 °C), while the presence of dopants in the poly-Si greatly enhances deuterium transport probability.

We use secondary ion mass spectrometry (SIMS) to characterize the hydrogen/deuterium distribution and concentration on 0.18 $\mu$m NMOS and PMOS poly-Si gate test structures subjected to deuterium anneal. Large area (500 $\mu$m×500 $\mu$m), planar gate stacks consisting of poly-Si gate/$SiO_2$ (4 nm)/Si(100) without nitride layers were examined by SIMS. Deuterated forming gas ($N_2$ 90%: $D_2$ 10%) was used as received without further purification. Wafers with the representative complementary MOS (CMOS) devices were pulled from the fabrication line and cleaved into pieces to enable placement in a tube furnace.

Dynamic SIMS sputter-depth-profiling analyses were carried out in on a Cameca IMS-6f magnetic sector instrument using an $O_2^+$ (8 keV) or $Cs^+$ (14.5 keV) ion beam raster scanned over a 125 $\mu$m×125 $\mu$m area on designated sample areas. Positive secondary ions ($Si^+$, $H^+$, $D^+$) were collected from the central 30-$\mu$m-diam area using $O_2^+$ ion sputtering as a function of sputter time (depth). Negative secondary ions ($Si^-$, $H^-$, $D^-$) were collected over a similar area using $Cs^+$ ion sputtering.

Simultaneous hydrogen/deuterium concentration quantification in Si is achieved by using a new, double-implanted (with H and D) standard in crystalline Si.[6] The secondary negative ion yield enhancement effect by Cs and the simultaneous suppression of molecular ion formation combine to provide significantly improved detection sensitivity for deuterium, at $\sim 10^{15}/cm^3$ (by 2-orders of magnitude over the positive ion mode) compared to our previous work.[4] The drawback, however, is a (well documented) complex behavior of negative secondary ion yield change across the $Si/SiO_2$ interface region, making *accurate* quantification of interfacial hydrogen/deuterium concentration difficult.

By contrast, the $Si/SiO_2$ interface profile in the positive secondary ion mode using $O_2^+$ sputtering is well behaved. A straightforward ion yield enhancement due to the presence of additional oxygen in the $SiO_2$ is observed. Thus in the calculation of interfacial Si–H(D) areal densities, we choose to use the positive secondary ion SIMS data, despite the rather poor overall positive secondary ion yield and molecular ion ($H_2^+$) limited deuterium detection background at m/e=2. In the poly-Si region, both the positive and negative secondary ion modes gave essentially identical SIMS profiles.

Figures 1 and 2 show, respectively, the positive and negative dynamic SIMS depth profiles obtained from the poly-$Si/SiO_2$ test block structures. Both NMOS and PMOS areas were examined after the source/drain (S/D) dopant activation anneal step (partially completed transistor flow) as a baseline reference and then after further $D_2$ sintering. The temperature employed in the S/D anneal (>900 °C in an inert $N_2$ ambient) is expected to result in significant *depassivation* of the $SiO_2$–Si interface (by Si–H bond scission) and *dehydrogenation* of the poly-Si. Similar thermal depopulation of hydrogen (as Si–H) from the $SiO_2$–Si interface and generation of $P_b$ centers have been widely studied by electron paramagnetic resonance (EPR) on vacuum annealed gate oxides at T>700 °C,[7] and by electrical characterization capacitance–voltage (C–V) of MOS capacitor structures with inert ambient rapid thermal anneal (RTA).[8]

The nature of the hydrogen present in the gate $SiO_2$ is believed to originate from thermally stable OH species throughout the gate $SiO_2$. The presence of small amount of such stable OH groups in thermally grown $SiO_2$ has been well documented in the literature.[9]

It is clear that following the 435 °C sinter in $N_2$:$D_2$ forming gas, Si–D enrichment is taking place in the gate oxide interface region. This observation is consistent with the *repassivation* of the interfacial dangling bonds Si· with

a)Electronic mail: rmwallace@ti.com

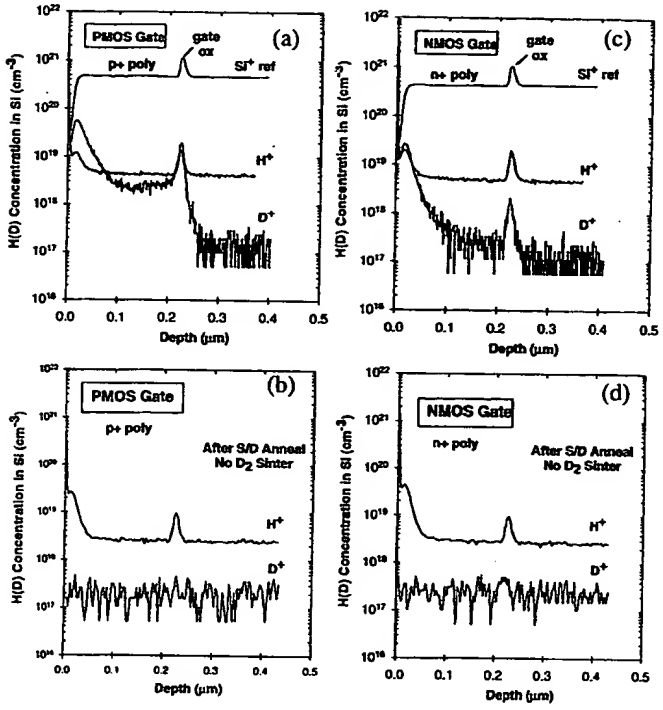


FIG. 1. Positive secondary ion SIMS profiles. (a) Doped $p^+$ poly-Si gate stacks after and before (b) $D_2:N_2$ sintering; (c) doped $n^+$ poly-Si gate stacks after and before (d) $D_2:N_2$ sintering.

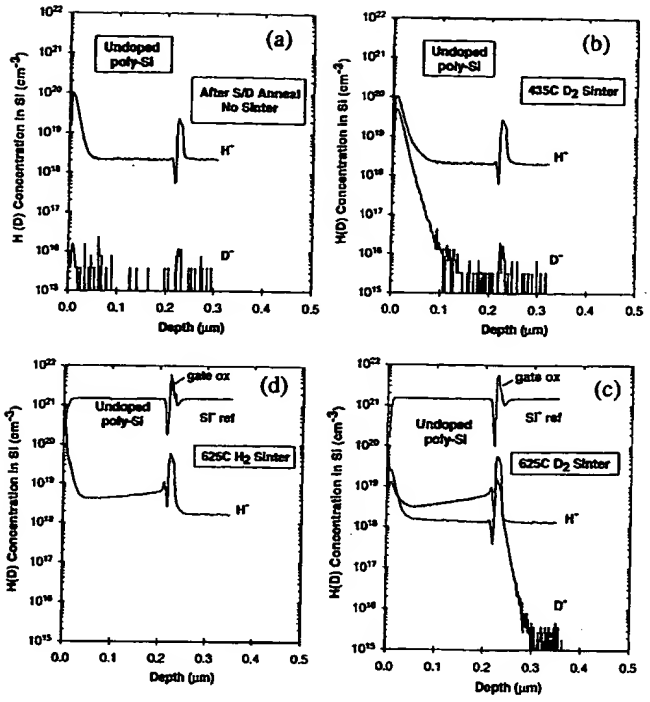


FIG. 3. Negative secondary ion ($M^-$) dynamic SIMS depth profiles obtained from undoped poly-Si on gate oxide stacks: (a) reference profile after S/D anneal; (b) $D_2$ sintered at 435 °C/1 h; (c) $D_2$ sintered at 625 °C/1 h; and (d) $H_2$ sintered at 625 °C/1 h.

deuterium (as Si–D).[4,5,8] This provides direct evidence that deuterium diffusion through planar poly-Si/$SiO_2$ structure can readily take place under the sintering conditions employed here, allowing efficient delivery of deuterium to the poly-Si/$SiO_2$/Si interfacial region.[2,5]

The quantified Si–D areal density values in the gate $SiO_2$–Si interface region are observed to be $2\times10^{12}/cm^2$ for NMOS and $1\times10^{13}/cm^2$ for PMOS.[10] The NMOS value is in close agreement with commonly accepted interfacial Si-dangling bond density ($D_{it}\sim1-2\times10^{12}$ $cm^{-2}$) used in hot carrier modeling, but the PMOS value is significantly higher, indicating additional deuterium exchange/incorporation has occurred. Indeed, the incorporation of deuterium into the bulk of the doped polysilicon is more significant in p+ poly (at $2-3\times10^{18}/cm^3$) than in n+ poly (at $\sim1\times10^{17}/cm^3$). This observation can be attributed to the formation of hydrogen-dopant complexes within the Boron doped poly-Si.[5,11]

Figure 3 presents the results of negative secondary ion SIMS depth profiles of un-doped polysilicon stack structures after various process conditions. Undoped poly-Si appears to be an effective barrier to deuterium diffusion upon sintering at 435 °C for 1 h [Fig. 3(b)]. In contrast to the D concentration observed after sintering in *doped* poly-Si stack structures (Fig. 2), the amount of D that can reach the gate oxide region is negligibly small in the undoped poly-Si stack.

However, if the $N_2:D_2$ sinter is conducted at a higher temperature, viz. 625 °C, then a significant increase in D concentration is once again observed in *both* the poly-Si itself *and* in the interface regions [Fig. 3(c)]. Similar sintering in $N_2:H_2$ at 625 °C [Fig. 3(d)] also leads to *rehydrogenation* of the undoped poly-Si *and* the interface regions.

These results from undoped poly-Si are very illuminating in that they clearly demonstrate the origin of molecular hydrogen "permeability" of the poly-Si material and further help to elucidate the hydrogen transport mechanism under hydrogen-containing ambient and sintering conditions in a typical CMOS device process flow.

In as-deposited poly-Si with low dopant concentration ($10^{15}-10^{16}$ $cm^3$), hydrogen (as Si–H) is believed to terminate and passivate the poly-Si grain boundaries.[12] Following a high temperature S/D anneal, *dehydrogenation* occurs and the Si-hydride (deuteride) is depleted [Fig. 3(a)]. Evidently, at 625 °C, *rehydrogenation* of the undoped poly-Si can readily occur and Si-hydride (deuteride) reforms presumably by reaction with *molecular hydrogen* ($H_2$) or deuterium ($D_2$). In such a case, transport of hydrogen (deuterium) is mediated by the formation of Si–H(D) species. This mechanism appears to be kinetically inhibited at 435 °C, rendering such sintering ineffective in undoped poly-Si.

The diffusion of hydrogen in crystalline silicon is a subject of significant complexity because hydrogen can be

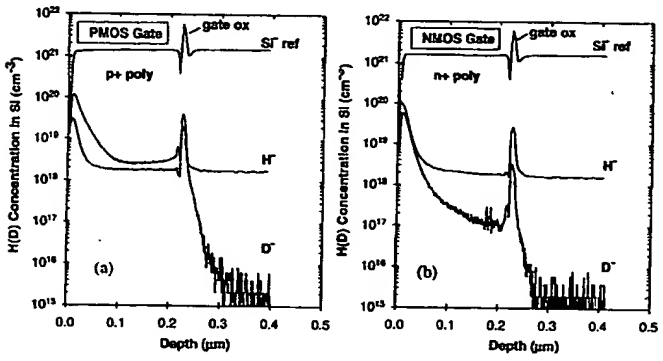


FIG. 2. Negative secondary ion SIMS profiles of doped (a) $p+$ and (b) $n$

with impurities (dopants) or defects.[11] Based on these published studies, the interstitial diffusion of *atomic hydrogen* ($H^0$) in monocrystalline Si is extremely facile at our sintering temperatures (400–600 °C). Since the observed deuterium interface passivation effects in doped poly-Si originates from *molecular deuterium* ($D_2$), its transport through poly-Si to the $SiO_2$–Si interface will *necessarily* involve a *dissociation* step, which is then most likely the rate-limiting step. Presence of dopants (B, P, As) in the poly-Si profoundly modify the formation mechanism of the hydrogenous species ($H^0$, $H^{\pm}$) and significantly lowers the *apparent* activation energy of the poly-Si *rehydrogenation* process. With increased hydrogen solubility in doped poly-Si, more efficient transport of hydrogen occurs at a significantly lower temperature (i.e., 435 °C).

The following elementary interfacial reactions have been proposed in the kinetic model for the passivation of the $SiO_2$–Si interface, and a steady-state equilibrium is assumed to be reached at the following sintering temperature:[7,13]

$$\mathrm{Si}\cdot + \mathrm{H} \rightarrow \mathrm{Si{-}H}, \tag{1}$$

$$\mathrm{Si{-}H} + \mathrm{H} \rightarrow \mathrm{Si}\cdot + \mathrm{H_2}. \tag{2}$$

Both Eqs. (1) and (2) assume supply of *atomic* hydrogen ($H^0$) or protonic ($H^{\pm}$) species to the $SiO_2$–Si interface is available during sintering.

In B-doped *monocrystalline* Si, the formation of a boron-hydrogen complex, designated as Si:B(H) has been well documented by both theoretical and numerous experimental observations,[11] including SIMS. This provides direct evidence that hydrogen transport will occur within individual crystalline Si grains. The role of grain boundary diffusion in heavily doped $p^+$ poly-Si is unclear at present, but if active, it is expected to further enhance hydrogen transport to the poly-Si/gate $SiO_2$ interface.

We also note that the Si:B(H) complex formation, well known to result in electrical deactivtion,[11] could be partially responsible for poly-Si depletion effects observed in PMOS devices. Recent reports utilizing poly $Si_xGe_{1-x}$ gates have indicated a higher fraction of B dopant activation.[14] With the reasonable assumption that similar Ge:B(H) formation and deactivation occurs, the subsequent enhanced reactivation efficiency at relatively low temperatures (~500 °C) could be consistent with a decreased H bond strength associated with nearby Ge compared to Si.[15] Finally, the role of B-impurity interactions and vacancies on electrical activation in *c*-Si has been recently reported and is not considered here.[16]

In $n^+$ poly-Si, phosphorous is known to segregate and *chemically passivate* the grain boundaries, resulting in a suppression of Si–H formation at grain boundaries. Nevertheless, upon sintering at 435 °C, significant transport of hydrogen (deuterium) can still take place as judged by the final increase in Si–D population at the $SiO_2$–Si interface, although much less D is found to incorporate into the $n^+$ poly-Si itself. The donor-hydrogen complex has also been reported in the literature for n-type dopants (P, As, and Sb) in *monocrystalline* Si and is once again the most likely intermediate responsible for hydrogen transport through the $n^+$ poly-Si.[11,12]

It is important to recognize this interface passivation poly-Si gate stack, particularly for CMOS structure with nitride sidewalls. The observed channel-hot-carrier lifetime improvement in NMOS devices by deuterium sintering[2–4] lends direct support to our observation of efficient deuterium transport through poly-Si, as the nitride sidewall is thought to be an effective barrier to deuterium diffusion.[2,3] Transport through oxide-sidewall CMOS structures is much different, where it is known that *molecular* hydrogen ($H_2$) can effectively transport through the $SiO_2$ network at low temperatures (225 °C and up) and potentially dominate the interface passivation process.[1,2,17]

Our present work also demonstrates that partially processed CMOS transistors are quite permeable by molecular or atomic hydrogen. In a typical CMOS processing flow, there exists a large number of temperature excursions and exposures to hydrogen-containing ambients, including encapsulation layers, all of which will affect the dehydrogenation/rehydrogenation reaction to differing degrees in both the poly-Si and the gate oxide interfaces.

The authors wish to thank Norman James for performing the $D_2$ sintering of samples. We are also grateful to Jiejie Xu for providing the CMOS test structures used in this study. Mark Anthony is gratefully acknowledged for useful discussions. The dynamic SIMS laboratory support of the Advanced Materials Characterization group is greatly appreciated.

[1] S. Wolf, in *Silicon Processing for the VLSI Era* (Lattice, Sunset Beach, CA, 1995), Vol. 3, p. 425.
[2] I. Kizilyalli, J. W. Lyding, and K. Hess, IEEE Electron Device Lett. **18**, 81 (1997).
[3] J. W. Lyding, K. Hess, and I. C. Kizilyalli, Appl. Phys. Lett. **68**, 2526 (1996).
[4] H. C. Mogul, L. Cong, R. M. Wallace, P. J. Chen, T. A. Rost, K. Harvey, and N. A. Dumin, Appl. Phys. Lett. **72**, 1721 (1998).
[5] P. J. Chen and R. M. Wallace, Mater. Res. Soc. Symp. Proc. (to be published).
[6] (a) H-implant in Si: standard reference material, Charles Evans & Associates, Redwood City, CA; (b) H and D implanted in Si by IICO Corp., CA. The quoted hydrogen dose was cross calibrated with standard (a) and was found to be within 5%. The deuterium dose was used as quoted.
[7] (a)J. H. Stathis, in *The Physics and Chemistry of $SiO_2$ and the Si–$SiO_2$ Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 393; (b)A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter **9**, 3299 (1997) and references therein; (c)A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter **10**, L19 (1998).
[8] P. K. Hurley, C. Leveugle, A. Mathewson, D. Doyle, S. Whiston, J. Prendergast, and P. Lungren, Mater. Res. Soc. Symp. Proc. (to be published).
[9] E. H. Poindexter, Z. Naturforsch. A **50**, 653 (1995).
[10] We estimate the repeatability of these measurements is ≤15%. The accuracy is estimated to be within a factor of 2, based on calibration of H and D in Si and applied to $SiO_2$ accounting for documented differences in relative ion yields in Si vs $SiO_2$.
[11] S. J. Pearton, J. W. Corbett, and M. Stavola, *Hydrogen in Crystalline Semiconductors* (Springer, New York, 1992), and collection of comprehensive references therein.
[12] J. Chevallier and M. Aucouturier, Annu. Rev. Mater. Sci. **18**, 219 (1988).
[13] K. L. Brower and S. M. Myers, Appl. Phys. Lett. **57**, 162 (1990).
[14] T.-J. King, J. P. McVittie, K. C. Saraswat, and J. R. Pfiester, IEEE Trans. Electron Devices **41**, 228 (1994).
[15] H. Kim, N. Taylor, J. R. Abelson, and J. E. Greene, J. Appl. Phys. **82**, 6062 (1994).
[16] F. Priolo, G. Mannino, M. Micciché, V. Privitera, E. Napolitani, and A. Carner, Appl. Phys. Lett. **72**, 3011 (1998).
[17] K. L. Brower, Appl. Phys. Lett. **53**, 508 (1998); Phys. Rev. B **38**, 9657

# Deuterium transport through device structure



P. J. Chen and R. M. Wallace[a)]
*Texas Instruments Incorporated, Dallas, Texas 75243*

(Received 19 January 1999; accepted for publication 17 May 1999)

We use secondary ion mass spectrometry to characterize the hydrogen/deuterium distribution and concentration on complimentary "metal" oxide silicon (CMOS) test structures subjected to molecular deuterium ($D_2$) anneals. We examine the temperature dependence and the influence of doping on the transport of deuterium to the gate oxide interfaces resulting in interface passivation. We find that *undoped polycrystalline silicon* appears to be an efficient barrier for deuterium transport at typical postmetallization sintering temperatures. We also examine the permeability of device structures that include dielectric encapsulation layers after typical postmetal sintering conditions employed in a conventional CMOS process flow. It is found that typical low temperature deposited oxide dielectrics are quite permeable by molecular deuterium at typical sintering temperatures (435 °C). In contrast, chemical vapor deposited silicon nitride dielectrics appear to form a complete barrier to deuterium diffusion (even for layers as thin as 300 Å). We also find that nitrides which receive a high thermal budget exposure, such as the source/drain anneal, appears to regain permeability to deuterium diffusion/transport. © *1999 American Institute of Physics.* [S0021-8979(99)07916-5]

## I. INTRODUCTION

Passivation of $SiO_2$–Si interface states through a "back end" process with molecular hydrogen ($H_2$) anneals (sinters) of metal oxide semiconductor field effect transistor (MOSFET) devices is a common and critical step in obtaining reliable high performance.[1] Additionally, reports have shown that *n*-type MOS (NMOS) channel hot carrier (CHC) lifetime can be markedly improved (10–50×) by low temperature (400–450 °C), post metallization (metal-1) sinter in molecular deuterium ($D_2$).[2–5] More recently, effects of $D_2$ sintering of multilevel metal structures with the associated interlevel dielectrics have also been reported.[6] In these processes, the efficient transport and delivery of such hydrogenous species to the $SiO_2$/Si interface is important to obtain the desired improvement. In the vicinity of the transistor region of an integrated circuit, a variety of potential transport pathways to the underlying gate dielectric interfaces exists, as depicted in Fig. 1. Species diffusing may encounter substantial barriers to diffusion through, for example, sidewall materials such as silicon nitride compared to silicon oxide.

In this work, we examine the transport of deuterium through various transistor gate stack layers to the gate dielectric interfaces arising from a $D_2$ sintering. We find that undoped poly-Si forms an efficient barrier to deuterium transport at typical sintering temperatures (~400–480 °C), while the presence of dopants in the poly-Si greatly enhances deuterium transport probability. We also find that the permeability of Si-nitride dielectric encapsulation layers is sensitive to prior thermal treatments in processing.

[a)]Present address: Materials Science Department, University of North Texas, Denton, Texas 76203-5310. Electronic mail: rwallace@unt.edu

## II. EXPERIMENT

We use dynamic secondary ion mass spectrometry (SIMS) to characterize the hydrogen/deuterium distribution and concentration on 0.18 $\mu$m NMOS and PMOS poly-Si gate test structures subjected to deuterium anneal. Large area (500 $\mu$m×500 $\mu$m), planar gate stacks with and without various dielectric encapsulation layers were examined by SIMS. Deuterated forming gas ($N_2$ 90%; $D_2$ 10%) was used as received without further purification. Wafers with the representative CMOS devices were pulled from the fabrication line and cleaved into pieces to enable placement in a tube furnace.

Dynamic SIMS sputter-depth-profiling analyses were carried out on a Cameca IMS-6f magnetic sector instrument using an $O_2^+$ (8 keV) or $Cs^+$ (14.5 keV) ion beam raster-scanned over a 125 $\mu$m×125 $\mu$m area on designated sample areas. Rigorous vacuum pump-down procedures were followed after each sample introduction to achieve best possible base pressure recovery (in the $10^{-10}$ Torr range) in order to minimize hydrogen background from the residual $H_2O$ and $H_2$ content. Positive secondary ions ($Si^+$, $H^+$, $D^+$) were collected from the central 30 $\mu$m diameter area using $O_2^+$ ion sputtering as a function of sputter time (depth). Negative secondary ions ($Si^-$, $H^-$, $D^-$) were collected over a similar area using $Cs^+$ ion sputtering.

Simultaneous hydrogen/deuterium concentration quantification in Si is achieved by using a new, double-implanted (with H and D) standard in crystalline Si.[7] The secondary negative ion yield enhancement effect by Cs and the simultaneous suppression of molecular ion formation combine to provide significantly improved detection sensitivity for deuterium, at ~$10^{15}$/cm$^3$ (by two orders-of-magnitude over the positive ion mode) compared to our previous work.[4] The drawback, however, is a (well-documented) complex behav-

 

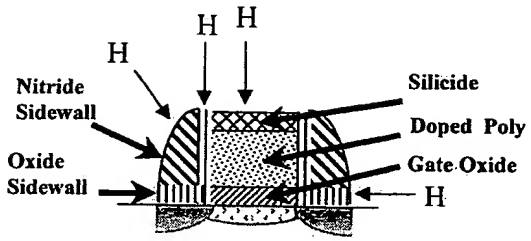


FIG. 1. Diffusion pathways for hydrogenous species in the vicinity of a transistor.

ior of negative secondary ion yield change across the $Si/SiO_2$ interface region, making *accurate* quantification of interfacial hydrogen/deuterium concentration difficult.

By contrast, the $Si/SiO_2$ interface profile in the positive secondary ion mode using $O_2^+$ sputtering is well-behaved. A straightforward ion yield enhancement due to the presence of additional oxygen in the $SiO_2$ is observed. Thus in the calculation of interfacial Si–H(D) areal densities, we choose to use the positive secondary ion SIMS data, despite the rather poor overall positive secondary ion yield and molecular ion ($H_2^+$) limited deuterium detection background at $m/e=2$. In the poly-Si region, both the positive and negative secondary ion modes gave essentially identical SIMS profiles.

## III. RESULTS

### A. Deuterium transport through poly-Si transistor gate stack

Three types of poly-Si-on-gate-$SiO_2$ stack were used for this study. The PMOS and NMOS gate stacks received standard CMOS transistor flow processing up to the source/drain (S/D) anneal step. The undoped poly-Si gate stack went through nearly identical processing steps (including S/D anneal), but without receiving any ion implant doping in the poly-Si gate. The temperature employed in the S/D anneal process (>900 °C in an inert $N_2$ ambient) is expected to result in significant *depassivation* of the $SiO_2$–Si interface (via Si–H bond scission) as well as substantial *dehydrogenation* in the poly-Si itself.

Similar thermal depopulation of hydrogen (conversion of Si–H to a Si dangling bond, denoted as Si•) at the $SiO_2$–Si interface and generation of trap states (so-called "$P_b$ centers") have been widely studied by electron paramagnetic resonance (EPR) on vacuum annealed gate oxides at $T>700$ °C,[8] and by electrical characterization *(C–V)* of MOS capacitor structures with inert ambient rapid thermal anneal (RTA).[9] The nature of the hydrogen present in the gate $SiO_2$ is believed to originate from thermally stable OH species throughout the gate $SiO_2$. The presence of a small amount of

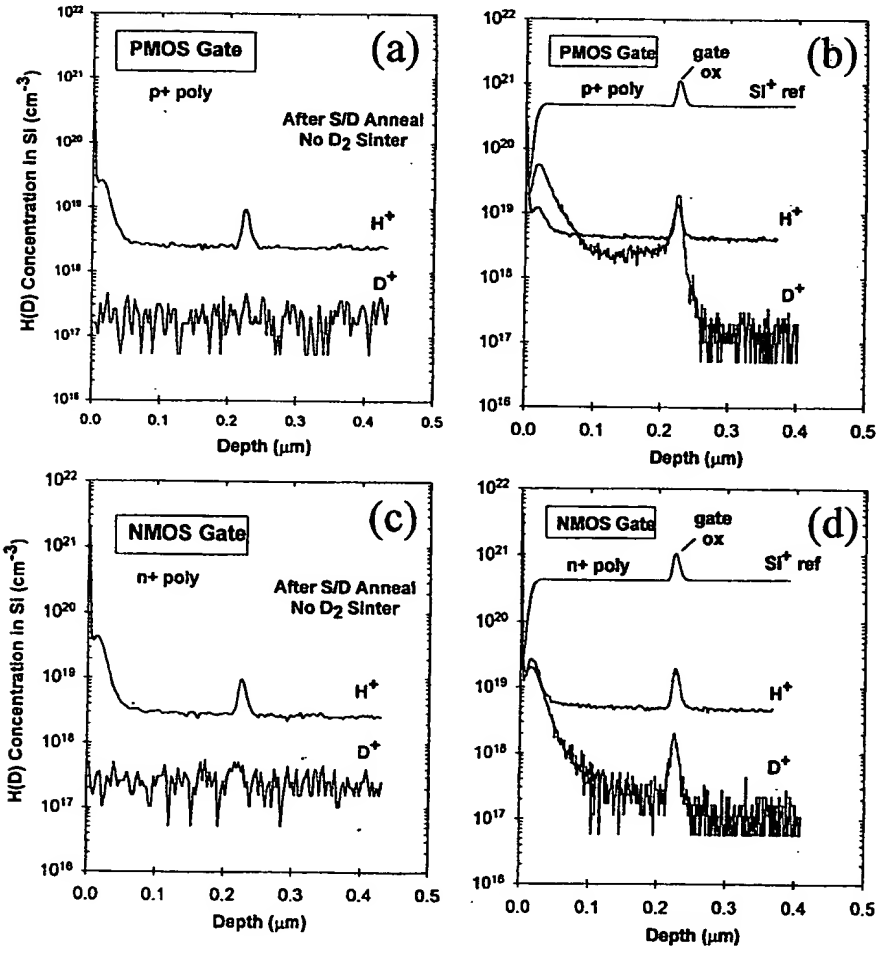


FIG. 2. Positive secondary ion SIMS profiles. (a) Doped $p^+$ poly-Si gate stacks before and after (b) $D_2$:$N_2$ sintering; (c) doped $n^+$ poly-Si gate stacks before and after (d) $D_2$:$N_2$ sintering.

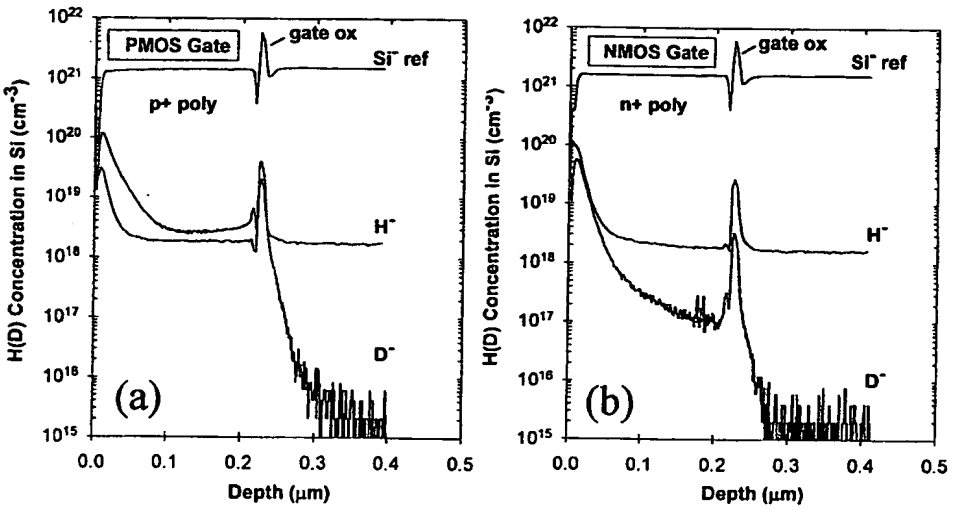


FIG. 3. Negative secondary ion SIMS profiles of doped (a) $p^+$ and (b) $n^+$ poly Si gate stacks after $D_2$:$N_2$ sintering at 435 °C for 1 h.

such stable OH groups in thermally grown $SiO_2$ has been well documented in the literature.[10]

Figures 2 and 3 show, respectively, the positive and negative dynamic SIMS depth profiles obtained from poly-Si/$SiO_2$/Si(100) test block structures. Both PMOS and NMOS areas were examined after the S/D dopant activation anneal step as a baseline reference [Figs. 2(a) and 2(c)] and then after $D_2$ sintering at 435 °C for 1 h [Figs. 2(b), 2(d), 3(a), and 3(b)]. The lower limit of H concentration levels observed in these SIMS depth profiles of poly-Si gate stack structures were determined by residual-gas background hydrogen ions produced in the SIMS apparatus during ion beam sputtering. The enhanced H concentration in the vicinity of the gate $SiO_2$ is attributed to the presence of stable OH groups as discussed previously.[10]

It is clear that following the 435 °C sinter in $N_2$:$D_2$ forming gas, Si–D enrichment is taking place in the gate oxide interface region as seen in Figs. 2(b), 2(d), 3(a), and 3(b). This observation is consistent with the *repassivation* of the interfacial dangling bonds Si• with deuterium (as Si–D).[4,9,11,12] This provides direct evidence that deuterium diffusion through planar poly-Si/$SiO_2$ structure can readily take place under the sintering conditions employed here, allowing efficient delivery of deuterium to the poly-Si/$SiO_2$/Si interfacial region.[11–13]

From the positive secondary ion SIMS data, the quantified Si–D areal density values in the gate $SiO_2$–Si interface region are observed to be $2\times10^{12}/cm^2$ for NMOS and $1\times10^{13}/cm^2$ for PMOS.[14] The NMOS value is in close agreement with commonly accepted interfacial Si-dangling bond density ($D_{it}\sim1-2\times10^{12}/cm^2$) used in hot carrier modeling, but the PMOS value is significantly higher, indicating additional deuterium exchange/incorporation has occurred.

Indeed, the incorporation of deuterium into the bulk of the doped polysilicon is more significant in $p^+$ poly (at $2-3\times10^{18}/cm^3$) than in $n^+$ poly (at $\sim1\times10^{17}/cm^3$), according to either the positive [Figs. 2(b) and 2(d)] or negative [Figs. 3(a) and 3(b)] secondary ion SIMS profiles. The observation of enhanced deuterium concentrations in PMOS versus NMOS can be attributed to the formation of hydrogen-dopant complexes within the poly-Si and will be discussed below.[11,12,15] We have also previously reported the observation of deuterium transport to the gate dielectric interface in devices with gate metal contacts on the poly-Si gate electrode as well.[4,12] It was found that metallized contact (titanium silicide) on poly-Si does not form a barrier to deuterium diffusion.

Figure 4 presents the results of negative secondary ion SIMS depth profiles of *undoped* polysilicon stack structures before [Fig. 4(a)] and after [Fig. 4(b)] identical sinter process conditions. Undoped poly-Si appears to be an effective barrier to deuterium diffusion upon sintering at 435 °C for 1 h [Fig. 4(b)]. In contrast to the D concentration observed after sintering in *doped* poly-Si stack structures (Fig. 3), the amount of D that can reach the gate $SiO_2$ region is negligibly small in the undoped poly-Si stack. Incorporation of D into the bulk of undoped poly-Si is also negligible, except for a rapid decaying surface tail distribution.

Examination of the temperature dependence of this diffusion process is shown in Fig. 4(c) where it is observed that if the $N_2$:$D_2$ sinter is conducted at a higher temperature, viz. 625 °C for 1 h, then a significant increase in D concentration is once again observed in *both* the poly-Si itself *and* in the interface region [Fig. 4(c)]. We note that the observed H concentration in Fig. 4(c) remains at the same level as that for Figs. 4(a) and 4(b) as this signal originates from the residual gas background and the stable OH species in the gate $SiO_2$ region described previously. Similar sintering in $N_2$:$H_2$ at 625 °C [Fig. 4(d)] also leads to rehydrogenation of the undoped poly-Si and the $SiO_2$ interface regions as can be seen by the enhanced H concentration in these regions.

These results from undoped poly-Si clearly demonstrate that the origin of molecular hydrogen "permeability" of the poly-Si material is related to the ability to form stable, hydride-like species in the poly-Si. Evidently, such hydrogen permeability can also be altered by the presence of dopants in the poly-Si—in both the PMOS and NMOS cases, doping enables hydrogen transport in poly-Si at a much lower sintering temperature (i.e., 435 °C) compared to the undoped poly-Si case.

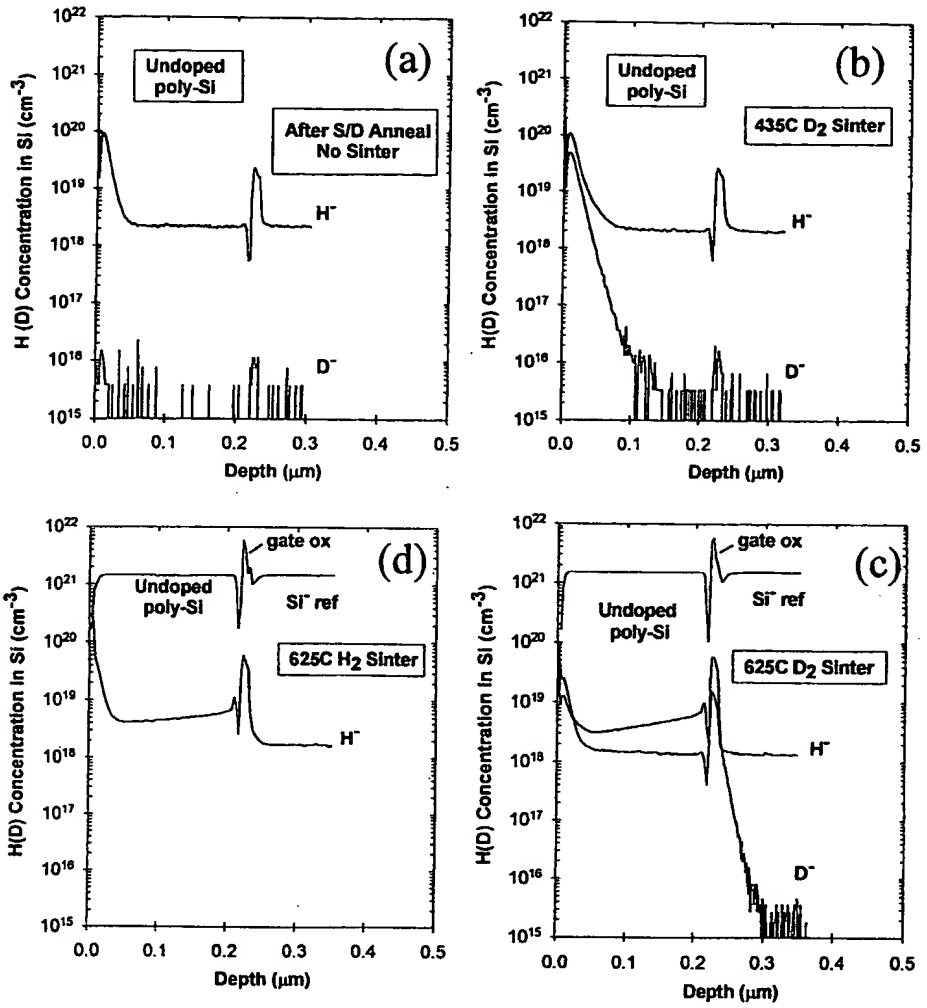


FIG. 4. Negative secondary ion ($M^-$) dynamic SIMS depth profiles obtained from undoped poly Si on gate oxide stacks: (a) reference profile after S/D anneal; (b) $D_2$ sintered at 435 °C/1 h; (c) $D_2$ sintered at 625 °C/1 h; and (d) $H_2$ sintered at 625 °C/1 h.

## B. Deuterium transport through Si–oxide/nitride dielectric encapsulation layers

Figure 5 shows the positive secondary ion ($M^+$) SIMS depth profiles obtained from a dielectric encapsulated CMOS device No. 1 before ($H_2^+$, $D^+$ ref) and after ($D^+$) deuterium ($D_2$) sintering. This device contains a thin film stack of borophosphosilicate glass (BPSG) and silicon oxide produced by a plasma-enhanced tetraethyloxysiloxane (PETEOS) deposition process resulting in the layers shown in the figure. Note that this stack has a LPCVD nitride encapsulation layer above the transistor gate stack. This device has undergone complete transistor flow and metal-1 process. An open test area of ~200 μm×280 μm was used for SIMS measurement after deuterium sinter at 435 °C/1 h.

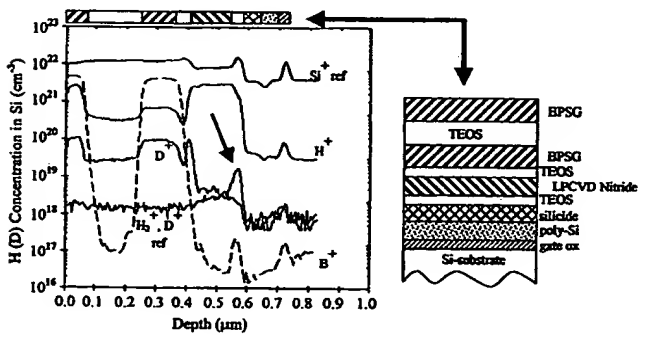


FIG. 5. Positive secondary ion ($M^+$) SIMS profiles obtained from $D_2$:$N_2$ sintered postmetal-1 CMOS device dielectric stack. Deuterium enrichment of at least 20–100× can be observed in both the BPSG and TEOS $SiO_2$ dielectric after the 435 °C/1 h $D_2$:$N_2$ sinter. Deuterium incorporation into the thin TEOS $SiO_2$ layer *underneath* the nitride at a depth of ~5500 Å is also observed. The lateral area of the planar gate stack is ~200 μm×280 μm.

From the SIMS elemental depth profiles in Fig. 5, it can be seen that deuterium incorporation into the PETEOS and BPSG layers is readily observed in the 435 °C/1 h $D_2$:$N_2$ sintered device. The estimated D concentration in the dielectric is on the order of $10^{19}$–$10^{20}$ $cm^{-3}$, based on documented SIMS sensitivity of D in $SiO_2$ and calibration of D in Si. More deuterium is found to incorporate in BPSG than in PETEOS $SiO_2$ (by ~2×). Deuterium incorporation into the thin TEOS Si-oxide layer *underneath* the planar LPCVD nitride (which served as sidewall spacer in the device structure) is also observed (as indicated by arrow in Fig. 5) at a lower concentration level. The boron observed in the gate oxide and beyond originates from the *p*-well doping in the Si substrate as the poly-Si gate for this NMOS device is P-doped.

The effect of an insertion of a PECVD nitride layer between the BPSG and TEOS layers is shown in Fig. 6. This

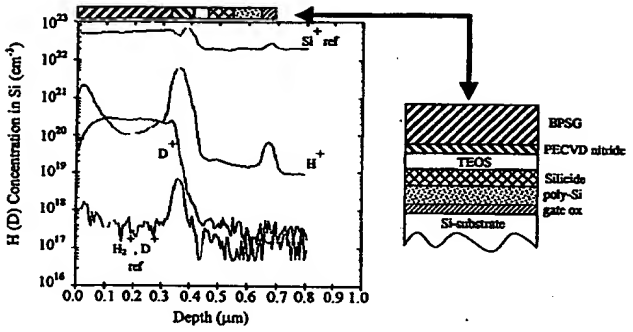


FIG. 6. Positive secondary ion ($M^+$) SIMS profiles obtained from $D_2$:$N_2$ sintered postmetal-1 CMOS device dielectric stacks. Deuterium enrichment of at least 200× can be observed in the densified BPSG dielectric layer after 435 °C/1 h $D_2$:$N_2$ sinter. A thin (~30 nm) PECVD nitride layer is underneath the BPSG and is the source of the strong hydrogen ($H^+$) signal.

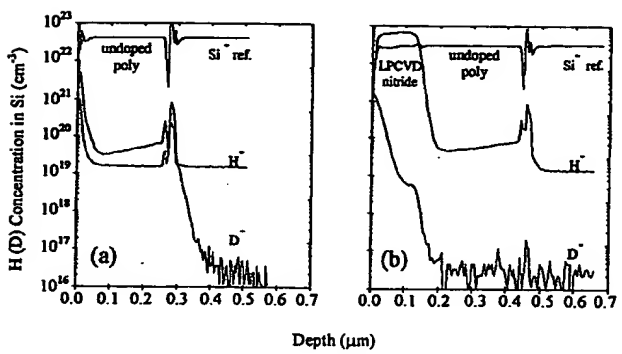


FIG. 7. Negative secondary ion ($M^-$) SIMS profiles obtained from (a) $D_2$ sintered poly-Si/gate oxide (10 nm) stack at 625 °C/1 h; and (b) LPCVD nitride/poly-Si/gate oxide stack. The as-deposited LPCVD nitride is found to form a complete barrier to deuterium transport (diffusion) at 625 °C.

device contains a thin film stack of BPSG/nitride/TEOS on top of the transistor gate stack before ($H_2^+$, $D^+$ ref.) and after ($D^+$) the device was sintered in deuterium after the metal-1 process step. The area used for SIMS analysis is identical to that used in poly-Si/gate stacks. The lateral area of the planar gate stack is ~450 μm×500 μm.

The SIMS results confirm the permeability of deuterium in the BPSG dielectric. However, the deposited PECVD nitride layer, despite being only 300 Å thick (at a depth of ~3500 Å), apparently forms a complete barrier to deuterium diffusion.

### C. Deuterium transport through silicon nitride dielectric encapsulation layers

The deposited silicon nitride film is generally regarded as a good barrier layer against moisture and atmospheric ambient permeation and is used almost exclusively as the last capping layer (protective overcoat) in CMOS backend process. A typical low temperature post metal anneal in forming gas is conducted to passivate interfacial states before (and occasionally after) the protective overcoat. Another significant role for silicon nitride films in CMOS is to serve as the sidewall spacer in the transistor structure. For deep submicron CMOS technologies, sidewall spacer material accounts for an increasingly larger fraction of the transistor volume. Gate dielectric interface engineering studies have also employed nitride layers. The deposition process and materials properties of such nitride layers critically impact many CMOS integration issues, such as B out-diffusion from the poly-Si gate electrode.

The deuterium annealing results on nitride-containing BPSG/TEOS thin film stacks seemed to suggest the permeability properties of the silicon nitride depends strongly on its processing history. To further clarify the properties of silicon nitride encapsulation dielectrics exclusively, a series of films were prepared without the associated BPSG or TEOS layers on pilot wafers. These are then compared to nitride films in processed device stacks with the same nitride process chemistry, but which have been subjected to the high temperature S/D anneals.

As-deposited LPCVD nitride films (using a $SiH_2Cl_2$ and $NH_3$ chemistry) were obtained and submitted to a deuterium sinter at 625 °C for 1 h. The sample structure consists of a planar LPCVD nitride capping layer (~150 nm) deposited on an undoped poly-Si gate oxide stack. For comparison, a sample that is not capped by the nitride was also submitted to the same sinter process. The negative secondary ion SIMS depth profile for these films are shown in Fig. 7. In agreement with the depth profiles of the undoped poly-Si device structures described previously (Fig. 4), the undoped poly-Si is found to incorporate deuterium in both the poly-Si and gate oxide region [Fig. 7(a)]. By contrast, the same undoped poly-Si gate oxide stack capped with LPCVD nitride does not exhibit observable permeability to deuterium after identical annealing in $D_2$:$N_2$ at 625 °C/1 h [Fig. 7(b)], as can be seen from lack of any deuterium incorporation into the underlying poly-Si and/or gate $SiO_2$ interface regions. From these SIMS results, it is apparent that the as-deposited LPCVD silicon nitride contains a high level of hydrogen, and forms a complete barrier to deuterium diffusion at 625 °C or below.

The effect of a high temperature anneal on a LPCVD nitride layer (also used as CMOS sidewall spacer) is shown in Fig. 8 where, during processing, this nitride layer was exposed to the S/D anneal process at $T>900$ °C. Unlike the polysilicon film, where dehydrogenation can readily occur after S/D anneal, the hydrogen content in the nitride film remained very high (estimated to be in the 0.1–1% range by SIMS), indicating the hydrogen in the LPCVD nitride is much more strongly bound. Nevertheless, for such sidewall nitride films that went through the S/D anneal step, a significant permeability to deuterium is observed [Figs. 6 and 8(b)] upon deuterium sintering.

## IV. DISCUSSION

### A. Hydrogen interaction with poly-Si and $SiO_2$-Si interface passivation

In as-deposited poly-Si with low dopant concentration ($10^{15}$–$10^{16}$ cm$^{-3}$), hydrogen (as Si–H) is believed to terminate and passivate the poly-Si grain boundaries.[16] Following a high temperature S/D anneal, dehydrogenation occurs and the Si-hydride (deuteride) is depleted [Figs. 2(a) and 2(c)]. Evidently, at 625 °C, rehydrogenation of the undoped poly-Si can readily occur and Si-hydride (deuteride) reforms

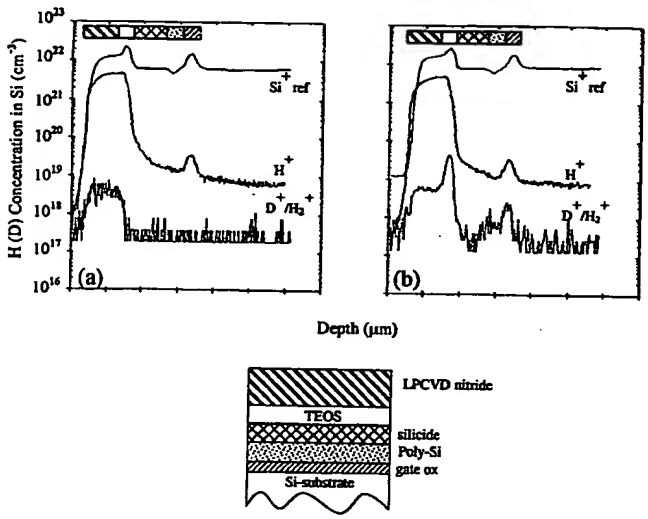


FIG. 8. Positive secondary ion ($M^+$) SIMS profiles obtained from LPCVD nitride/TEOS $SiO_2$/$WSi_x$/poly/gate oxide stack after (a) S/D anneal and (b) $D_2$ sinter at 435 °C/1 h. Deuterium is found to diffuse through the nitride layer (after S/D anneal) and become incorporated into the underlining TEOS $SiO_2$ and poly-Si/gate oxide region. The lateral area of the planar gate stack is ~200 $\mu$m×280 $\mu$m.

presumably by reaction with molecular hydrogen ($H_2$) or deuterium ($D_2$). In such a case, transport of hydrogen (deuterium) is mediated by the formation of Si–H(D) species. This mechanism appears to be kinetically inhibited at 435 °C, rendering such sintering ineffective in undoped poly-Si. We note that this thermal behavior is consistent with the observed Si–H(D) bond scission kinetics from Si surface thermal desorption studies.[17]

The diffusion of hydrogen in crystalline silicon is a subject of significant complexity because hydrogen can be present in several forms ($H_2$, $H^0$, $H^{\pm}$) and interact strongly with impurities (dopants) or defects.[15] Based on these published studies, the interstitial diffusion of atomic hydrogen ($H^0$) in monocrystalline Si is extremely facile at our sintering temperatures (400 °C–600 °C). Since the observed deuterium interface passivation effects in doped poly-Si originate from molecular desorption ($D_2$), its transport through poly-Si to the $SiO_2$–Si interface will necessarily involve a dissociation step, which is then most likely the rate-limiting step. Presence of dopants (B, P, As) in the poly-Si profoundly modify the formation mechanism of the hydrogenous species ($H^0$, $H^{\pm}$) and significantly lowers the apparent activation energy of the poly-Si rehydrogenation process. With increased hydrogen solubility in doped poly-Si, more efficient transport of hydrogen occurs at a significantly lower temperature (i.e., 435 °C).

The following elementary interfacial reactions have been proposed in the kinetic model for the passivation of the $SiO_2$–Si interface, and a steady-state equilibrium is assumed to be reached at the sintering temperature:[8,18]

$$Si\bullet + H \rightarrow Si\text{–}H, \tag{1}$$

$$Si\text{–}H + H \rightarrow Si\bullet + H_2. \tag{2}$$

Both Eqs. (1) and (2) assume supply of atomic hydrogen ($H^0$) or protonic ($H^{\pm}$) species to the $SiO_2$–Si interface is available during sintering.

In B-doped monocrystalline Si, the formation of a boron–hydrogen complex, designated as Si:B(H) has been well-documented by both theoretical and numerous experimental observations,[15] including SIMS. This provides direct evidence that hydrogen transport will occur within individual crystalline Si grains. The role of grain boundary diffusion in heavily doped $p^+$ poly-Si is unclear at present, but if active, it is expected to further enhance hydrogen transport to the poly-Si/gate $SiO_2$ interface.

We also note that the Si:B(H) complex formation, well known to result in electrical deactivation,[15] could be partially responsible for poly-Si depletion effects observed in PMOS devices. Recent reports utilizing poly $Si_xGe_{1-x}$ gates have indicated a higher fraction of B (and P) dopant activation for $0.25 \leqslant x \leqslant 0.45$.[19] With the reasonable assumption that similar Ge:B(H) formation and deactivation occurs, the subsequent enhanced reactivation efficiency at relatively low temperatures (~500 °C vs 900 °C) could be consistent with a decreased H bond strength associated with nearby Ge (Ref. 20) compared to Si.[17] If we use, for example, surface desorption studies of hydrogen as a trend to similar (bulk) bond dissociation energies, the literature clearly indicates a lower desorption peak maximum for $GeH_x$ (~250 °C) vs $SiH_x$ (~500 °C), and correspondingly reduced activation energies of desorption for Ge vs Si. Thus, assuming that the B(H) complex bond strength is similarly perturbed by the presence of Ge, a weaker bond to H could be anticipated. We also note that B–H bonding is expected to be weaker than that for Si–H as observed in surface desorption measurements[21] and from recent *ab initio* calculations comparing the bulk B–H vs various Si–H bond strengths.[22]

As a result of subsequent thermal treatments of the poly-SiGe film, one would therefore expect the passivating H concentration to be reduced relative to that in poly-Si as the hydrogen would be expected to diffuse more readily in the poly-SiGe to interfaces, etc. Finally, we note that the role of B-impurity interactions and vacancies on electrical activation in *c*-Si has been recently reported and is not considered here.[23]

In $n^+$ poly-Si, phosphorous is known to segregate and chemically passivate the grain boundaries, resulting in a suppression of Si–H formation at grain boundaries. Nevertheless, upon sintering at 435 °C, significant transport of hydrogen (deuterium) can still take place as judged by the final increase in Si–D population at the $SiO_2$–Si interface, although much less D is found to incorporate into the $n^+$ poly-Si itself. The donor-hydrogen complex has also been reported in the literature for *n*-type dopants (P, As, and Sb) in monocrystalline Si and is once again the most likely intermediate responsible for hydrogen transport through the $n^+$ poly-Si.[15,16]

It is important to recognize this interface passivation pathway by dopant-mediated hydrogen transport through the poly-Si gate stack, particularly for CMOS structure with nitride-sidewalls. The observed channel-hot-carrier lifetime improvement in NMOS devices by deuterium sintering[4,13]

lends direct support to our observation of efficient deuterium transport through poly-Si, as the nitride sidewall is thought to be an effective barrier to deuterium diffusion. Transport through oxide-sidewall CMOS structures is much different, where it is known that molecular hydrogen ($H_2$) can effectively transport through the $SiO_2$ network at low temperatures (225 °C and up) and potentially dominate the interface passivation process.[1,2,24]

### B. Hydrogen interaction with silicon nitride dielectric

Our results from deposited silicon nitride films suggest that residual H in the deposited (LPCVD, PECVD) silicon nitride films is bound in the form of N–H species, resulting in a much higher thermal stability than Si–H[17,25] due to the higher N–H bond dissociation energies.[26] These strongly bound H in the silicon nitride apparently do not participate in a dehydrogenation process at typical sintering temperatures. As a result, no deuterium incorporation or permeability was observed in these nitride films [Figs. 6 and 7(b)].

However, for LPCVD silicon nitride used as sidewall spacers (Fig. 1), the device process flow is such that the nitride film necessarily receives a much higher thermal budget, namely the S/D anneal at 900 °C or above (in an inert ambient). This high temperature processing history is believed to be responsible for modifying the properties of the LPCVD silicon nitride film with respect to deuterium transport. Our results (Figs. 5 and 8) indicated that these nitride films regain a certain degree of permeability to deuterium transport. The apparent lack of extensive D/H isotopic exchange in the nitride film suggests only a small fraction of the H species is participating and responsible for the observed permeability. Therefore, this reestablished permeability to deuterium diffusion (compared to no detected permeability in as-deposited LPCVD nitride) could possibly arise from dehydrogenation of a small fraction of more weakly bound H species in the LPCVD nitride film during S/D anneal. This dehydrogenation then enables reformation of weakly bonded hydride intermediates that lead to transport of deuterium at subsequent low temperatures used in the sintering process (i.e., 435 °C).

In the depth profile shown in Fig. 6, a thin (~300 Å) layer of PECVD nitride layer is present between the top BPSG and the PETEOS $SiO_2$ liner above the silicide. This low deposition temperature (~400 °C) nitride (PECVD) was deposited after the CMOS transistors have been completed. As a result, the nitride film was not exposed to a significant thermal budget such as the S/D anneal process. As can be clearly seen from the SIMS depth profiles in Fig. 6, this PECVD nitride layer once again forms a complete barrier to deuterium diffusion/transport, very similar to the behavior of as-deposited LPCVD nitride shown in Fig. 7(b).

Thus we conclude that the permeability properties of silicon nitride depends strongly on its prior (thermal) processing history. The high hydrogen content and thermal stability of N–H bond in the deposited silicon nitride films suggest that even the high temperature S/D anneal only affects a minor fraction of the (more weakly bonded) H species in the nitride, possibly causing selective, partial dehydrogenation. It is then sufficient to reestablish permeability to hydrogen (deuterium) transport in the nitride during subsequent low temperature sintering. We note that similar postdeposition annealing effects have been very recently reported for nitride/oxide gate dielectrics produced by plasma-enhanced methods.[27,28]

In regard to CHC lifetime improvement reports on devices with nitride sidewall spacers, one finds a range of results from "no significant" increase,[29] a "1.5–2×" increase,[13] and a ">10×" increase[4] in CHC lifetime for the same technology under sintering conditions similar to those reported here. Noting that it is important to carefully compare the metric used in extrapolating the CHC lifetime, we speculate that the detailed thermal processing history of the associated nitride layer in the structure, and therefore the D permeability, could be responsible for the reported range of electrical measurement results and the sinter "soak time" behavior.

## V. CONCLUSION

We have studied the transport of deuterium in a simple gate stack and dielectric-encapsulated planar CMOS structures under typical postmetal sinter conditions. The following conclusions can be drawn from SIMS results:

Low temperature isolation dielectrics such as doped silicate glass (BPSG) and oxide layers produced from TEOS appear to be readily permeable to deuterium diffusion at typical sintering temperatures.

The permeability of CVD nitrides appears to depend on their process (thermal budget) history. "As-deposited" LPCVD nitrides, which have substantial hydrogen concentrations, form a complete barrier to deuterium transport at 625 °C or below. Subsequent (high temperature) thermal processing (S/D anneal) of such LPCVD nitride films apparently reestablishes some deuterium transport through the film. We also note that the PECVD nitride film, when not exposed to a high temperature S/D process step, forms a similar barrier to deuterium diffusion as the LPCVD nitrides.

It is evident that, in the course of CMOS process flows, there exists a large number of temperature excursions and exposures to hydrogen-containing ambients, including encapsulation layers, all of which will affect the dehydrogenation/rehydrogenation reaction to differing degrees in both the poly-Si and the gate oxide interfaces. Optimization of the H concentration in these films is expected to dramatically effect the efficient delivery of H to the $Si/SiO_2$ gate dielectric interfaces.

## ACKNOWLEDGMENTS

The authors wish to thank Norman James for performing the $D_2$ sintering of samples. They are also grateful to Jiejie Xu for providing the CMOS test structures used in this study. We also thank Paul Tiner of the Diffusion/Implant Branch in KFAB for providing the blanket nitride films for this study. Mark Anthony is gratefully acknowledged for useful discussions. The dynamic SIMS laboratory support of the Advanced Materials Characterization group is greatly appreciated.

[1] S. Wolf, in *Silicon Processing for the VLSI Era* (Lattice, Sunset Beach, CA, 1995), Vol. 3, p. 425.
[2] I. Kizilyalli, J. W. Lyding, and K. Hess, IEEE Electron Device Lett. **18**, 81 (1997).
[3] J. W. Lyding, K. Hess, and I. C. Kizilyalli, Appl. Phys. Lett. **68**, 2526 (1996).
[4] H. C. Mogul, L. Cong, R. M. Wallace, P. J. Chen, T. A. Rost, K. Harvey, and N. A. Dumin, Appl. Phys. Lett. **72**, 1721 (1998).
[5] Z. Chen, J. Lee, and J. Lyding, in *Hydrogen in Semiconductors and Metals*, edited by N. H. Nickel, W. B. Jackson, R. C. Bowman, and R. G. Liesure, Mater. Res. Soc. Symp. Proc. **513**, 313 (1998).
[6] I. C. Kizilyalli, G. Weber, Z. Chen, G. Abeln, M. Schofield, B. Kotzias, F. Register, E. Harris, S. Sen, S. Chetlur, M. Patel, L. Stirling, R. Huang, A. Massengale, P. K. Roy, G. Higashi, E. Foley, J. Lee, J. Lyding, and K. Hess, IEEE IEDM Tech. Digest, p. 935 (1998).
[7] (a) *H-implant in Si: Standard Reference Material* (Charles Evans & Associates, Redwood City, CA); (b) H- and D-implanted in Si by IICO Corp., CA. The quoted hydrogen dose was cross-calibrated with standard (a) and was found to be within 5%. The deuterium dose was used as-quoted.
[8] (a) J. H. Stathis, in *The Physics and Chemistry of $SiO_2$ and the Si-$SiO_2$ Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 393; (b) A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter **9**, 3299 (1997), and references therein, **10**, L19 (1998).
[9] P. K. Hurley, C. Leveugle, A. Mathewson, D. Doyle, S. Whiston, J. Prendergast, and P. Lungren (presented at the Spring MRS Symposium, 1998).
[10] E. H. Poindexter, Z. Naturforsch. A **50a**, 653 (1995).
[11] P. J. Chen and R. M. Wallace, Appl. Phys. Lett. **73**, 3441 (1998).
[12] P. J. Chen and R. M. Wallace, in *Hydrogen in Semiconductors and Metals*, edited by N. H. Nickel, W. B. Jackson, R. C. Bowman, and R. G. Liesure, Mater. Res. Soc. Symp. **513**, 325 (1998).
[13] J. Lee, S. Aur, R. Eklund, K. Hess, and J. W. Lyding, J. Vac. Sci. Technol. A **16**, 1762 (1998).
[14] We estimate the repeatability of these measurements is $\leq 15\%$. The accuracy is estimated to be within a factor of 2, based on calibration of H and D in Si and applied to $SiO_2$ accounting for documented differences in relative ion yields in Si vs $SiO_2$.
[15] S. J. Pearton, J. W. Corbett, and M. Stavola, in *Hydrogen in Crystalline Semiconductors* (Springer-Verlag, New York, 1992), and collection of comprehensive references therein.
[16] J. Chevallier and M. Aucouturier, Annu. Rev. Mater. Sci. **18**, 219 (1988).
[17] H. N. Waltenburg and J. T. Yates, Jr., Chem. Rev. **95**, 1589 (1995), and references therein.
[18] K. L. Brower and S. M. Myers, Appl. Phys. Lett. **57**, 162 (1990).
[19] T.-J. King, J. P. McVittie, K. C. Saraswat, and J. R. Pfiester, IEEE Trans. Electron Devices **41**, 228 (1994).
[20] H. Kim, N. Taylor, J. R. Abelson, and J. E. Greene, J. Appl. Phys. **82**, 6062 (1997).
[21] P. J. Chen, M. L. Colaianni, and J. T. Yates, Jr., J. Appl. Phys. **72**, 3155 (1992).
[22] C. G. Van de Walle, Mater. Res. Soc. Symp. Proc. **513**, 55 (1998).
[23] F. Priolo, G. Mannino, M. Micciché, V. Privitera, E. Napolitani, and A. Camera, Appl. Phys. Lett. **72**, 3011 (1998).
[24] K. L. Brower, Appl. Phys. Lett. **53**, 508 (1998); Phys. Rev. B **38**, 9657 (1988).
[25] R. Walsh, Acc. Chem. Res. **14**, 246 (1981).
[26] *CRC Handbook of Chemistry and Physics*, 67th ed., edited by R. C. Weast (CRC, Boca Raton, 1986).
[27] Y. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998).
[28] Y. Wu, G. Lucovsky, and H. Massoud, 1998 International Reliability Physics Symposium Proceedings 70.
[29] S. Aur, T. Grider, V. McNeil, T. Holloway, and R. Eklund, 1998 International Reliability Physics Symposium Proceedings 180.

# Deuterium sintering of silicon-on-insulator structures: D diffusion and replacement reactions at the $SiO_2$/Si interface

R. M. Wallace,[a] P. J. Chen, L. B. Archer, and J. M. Anthc
*Texas Instruments Incorporated, Dallas, Texas 75243*



(Received 18 February 1999; accepted 18 June 1999)

We use dynamic secondary ion mass spectrometry (SIMS) to examine the mechanism of H (D) incorporation into and retention within a buried $SiO_2$ film at 625 °C. We find that diffusion of $H_2$ ($D_2$) through the Si/$SiO_2$/Si structure at this temperature is facile and that isotopic exchange occurs at the interfaces upon subsequent forming gas anneals at 625 °C. A detailed examination of the isotopic exchange process indicates that the interfaces do not exhibit equivalent behavior. We also describe the artifacts observed in the SIMS profiles by comparing positive and negative secondary ion profiles. © *1999 American Vacuum Society.* [S0734-211X(99)01805-3]

## I. INTRODUCTION

The incorporation of a positive charge in buried $SiO_2$ films has been the subject of investigation for many years. It has, for example, been conclusively demonstrated using a variety of spectroscopic and electrical characterization methods that annealing of such films in a hydrogen ambient for $T \geqslant 400$ °C results in a positively charged layer thought to originate from an oxygen deficient region near the Si/$SiO_2$ interface.[1–4] The underlying physical mechanisms for the production of hydrogenic species under such annealing conditions have also been presented.[3,5]

More recent reports have shown that such $H_2$ annealing of $SiO_2$ films in forming gas can result in the production of mobile protons throughout the film.[6–8] The observed proton mobility can be used to generate a nonvolatile memory device by applying an appropriate bias across the oxide resulting in the formation of a "sheet" of positive charge near the $SiO_2$/Si interface. This positive charge sheet results in a substantial shift in the $I-V$ characteristics resulting in a reproducible hysteresis behavior that exhibits nonvolatile memory properties. It has also been reported that the density of mobile protons, and therefore the magnitude of the hysteresis, is significantly affected by the chemical kinetics of the $H_2$ annealing process.[9] Finally, the response of such mobile protons to irradiation has also been presented.[10,11]

The efficient transport of hydrogen to the interface region appears to be a key ingredient resulting in the observed phenomenon. All of the studies cited above use sensitive spectroscopic (e.g., electron paramagnetic resonance) or electrical (e.g., $C-V$ measurements) characterization as a means to study the observed phenomenon. Very recently, a study using $C-V$ measurements on area-delineated samples was conducted to better elucidate the transport of hydrogen to the Si/$SiO_2$ region where lateral transport through the $SiO_2$ layer was surmised.[12]

In an effort to correlate some aspects of the previous findings, we have performed direct physical characterization of annealed buried $SiO_2$ layers (produced by the Unibond® method) using dynamic secondary ion mass spectrometry (SIMS) techniques. Very early work by Gale *et al.* demonstrated the ability of SIMS to not only examine the spatial distribution of H in thick $SiO_2$ films, but to also directly observe the *redistribution* of H within such films upon electron injection measurements in metal-oxide-semiconductor (MOS) capacitor structures.[13] We note that, in contrast to our previous SIMS work on conventional complementary metal-oxide-semiconductor (CMOS) devices (with extremely thin gate dielectric layers),[14] the silicon-on-insulator (SOI) films reported here are ideally suited to retain superior depth resolution under SIMS investigation.

We first examine the complications associated with the acquisition of SIMS data from such insulating samples prone to charging effects. We then examine the incorporation of the D within the Si/$SiO_2$/Si structure after various thermal treatments. We also examine the propensity for Si–D formation from isotopic switching experiments.

## II. EXPERIMENT

Unibond® samples[15] were obtained and diced into smaller pieces ~1×1 cm. Anneals were performed in a 6 in. tube furnace with a clean quartz tube. Sample temperatures were calibrated prior to sintering with thermocouple measurements (thin gauge wire) using wafer quarters to insure accuracy for the gas flow conditions through the tube. A substantial deviation (~25 °C) between the furnace set point and the wafer surface temperature was observed at the sintering conditions employed. The temperatures reported herein are corrected for this deviation. All anneals were performed at the temperatures indicated in Table I followed by a rapid thermal quench (~100 °C/s) to room temperature by manually pulling the samples out of the furnace tube, unless noted.[9] A cross sectional schematic of the Unibond® SOI material is shown in Fig. 1, where a crystalline Si overlayer (200 nm) is placed upon a thermally grown $SiO_2$ (400 nm) layer on a Si substrate through a unique wafer bonding process.[16] Forming gas $N_2$:$H_2$ (90%:10%) used in the control sample was from site supply sources. The deuterated forming gas $N_2$:$D_2$ (90%:10%) was obtained from Spectra Gases and

[a]Present address: Dept. of Materials Science, Univ. of North Texas, Denton, TX 76203; electronic mail: rwallace@unt.edu

TABLE I. Samples used for deuterium sintering of Unibond®.

| Sample | Treatment | Temp./time |
|---|---|---|
| A-1 | As received | ... |
| A-2 | $N_2$:$H_2$ (90%:10%) | 625 °C/60 min |
| A-3 | $N_2$:$D_2$ (90%:10%) | 625 °C/60 min |
| A-4 | $N_2$:$D_2$ (90%:10%) | 435 °C/60 min |
| A-5 | "A-3"+$N_2$ | 125 °C/180 min |
| A-6 | "A-4"+$N_2$ | 185 °C/180 min |
| A-7 | "A-3"+$N_2$:$H_2$ | 625 °C/60 min |
| B-1 | $N_2$:$H_2$ | 625 °C/60 min |
| B-2 | "B-1"+$N_2$:$D_2$ | 625 °C/60 min |
| B-3 | "B-2"+$N_2$:$H_2$ | 625 °C/60 min |
| B-4 | "B-3"+$N_2$:$D_2$ | 625 °C/60 min |
| C-1 | $N_2$:$D_2$ | 625 °C/60 min |
| C-2 | "C-1"+$N_2$:$H_2$ | 625 °C/60 min |
| C-3 | "C-2"+$N_2$:$D_2$ | 625 °C/60 min |
| C-4 | "C-3"+$N_2$:$H_2$ | 625 °C/60 min |
| D-1 | $N_2$:$D_2$ | 625 °C/60 min<br>Slow cool |

used as received without further purification.[17] Anneals in pure $N_2$ were simply performed in a laboratory oven with flowing $N_2$.

Dynamic SIMS sputter-depth-profiling analyses were carried out using an $O_2^+$ (8 keV) or a $Cs^+$ (14.5 keV) ion beam raster-scanned over a 150 $\mu$m×150 $\mu$m area. Positive secondary ions ($Si^+$,$H^+$,$D^+$) were collected from the central 60-$\mu$m-diam area as a function of sputter time (depth). For positive secondary ion acquisition, an electron flood gun at 9 keV was also employed to compensate for sample charging. The need to employ an electron beam for sample charge compensation produces a higher $H^+$ ($m/e=1$) and $H_2^+$ ($m/e=2$) background level as a result of electron beam induced dissociation of vacuum residual gases (e.g., $H_2O$ dissociation to form $H^+$ and $H_2^+$ ions) as well as electron-stimulated desorption. Negative secondary ions ($Si^-$, $H^-$, $D^-$) were collected over a similar area using $Cs^+$ ion beam sputtering, along with an electron beam at near zero impact energy for charge compensation.

Atomic hydrogen/deuterium concentration quantification in Si is achieved by using a H-implanted standard.[18] Quantification was accomplished by employing a single relative sensitivity factor for both $m/e=1$ and $m/e=2$ data, normalized to a single Si count rate acquired at the end of the profile run. For the convenience of data presentation, the same sensitivity factor was used for hydrogen/deuterium in the Si and $SiO_2$ layers.

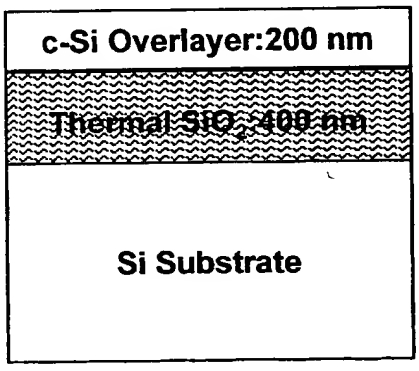


FIG. 1. Schematic of Unibond® SOI structure examined by dynamic SIMS.

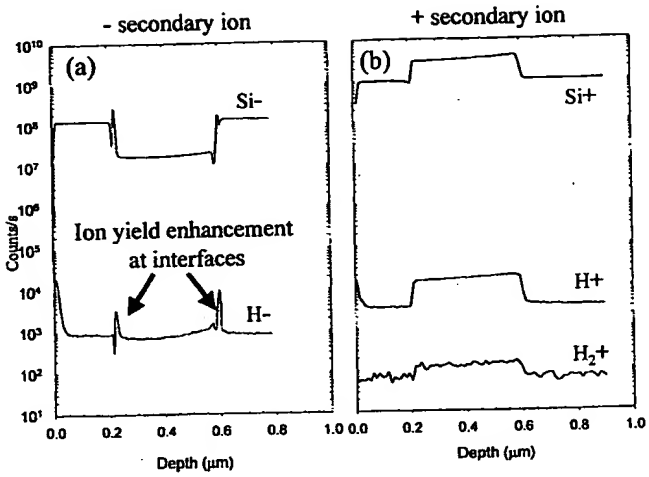


FIG. 2. Comparison of (a) negative and (b) positive secondary ion signals from SIMS analysis of as-received SOI material (sample A-1).

## III. RESULTS

### A. Complications in SIMS acquisition

We first examine the negative and positive secondary ion depth profiles shown in Fig. 2 from an "as-received" Unibond® sample (A-1). The SIMS profile clearly indicates the location of the $SiO_2$ layer and $Si/SiO_2$ interface by the abrupt change in secondary ion yield for $m/e$ 28 ($Si^+$, $Si^-$), $m/e$ 1 ($H^+$, $H^-$), and $m/e$ 2 ($H_2^+$).

Figure 2(a) presents the negative secondary ion depth profile where a significant ion yield enhancement is noted at the shallow and deep $Si/SiO_2$ interfaces. This effect, well documented in the SIMS literature, makes quantification of the interfacial hydrogen/deuterium difficult. For this reason, positive secondary ions are frequently chosen for quantitative interfacial concentration information.

Figure 2(b) presents such a positive secondary ion profile where a well-behaved ion yield transition across the interface is observed. We note that, in either profile, the H-ion yield throughout the $SiO_2$ layer corresponds to the lower limit of detection of such species as these as-received films are expected to exhibit a very low H content due to the extremely high temperatures employed in the Unibond® SOI growth process (described below). The detailed structure of the $SiO_2$/Si interface in the as-received sample is not known, but Si dangling bonds are observed to exist based on the sample fabrication procedure.[19,20]

We now consider the SIMS profiles of as-received samples that have been intentionally implanted with protons ($H^+$) shown in Fig. 3. Figure 3(a) presents the negative secondary ion profile from an implanted sample, with the $H^+$ implant energy (55 keV) selected so as to provide a peak concentration (total dose $=3\times10^{15}/cm^2$) near the deep $Si/SiO_2$ interface. The resultant (expected) asymmetric im-

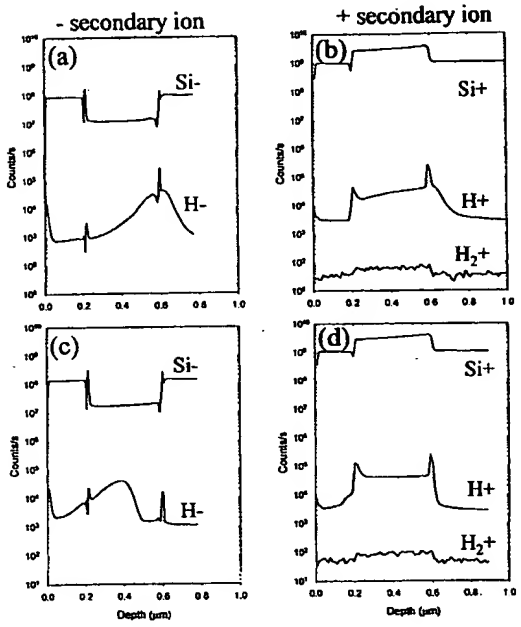


FIG. 3. SIMS depth profiles for H implanted SOI structure. (a) Negative secondary ion profile, implant peak at deep $SiO_2$/Si interface, (b) corresponding positive secondary ion profile, (c) negative secondary ion profile, implant at center of buried $SiO_2$ layer, (d) corresponding positive secondary ion profile.

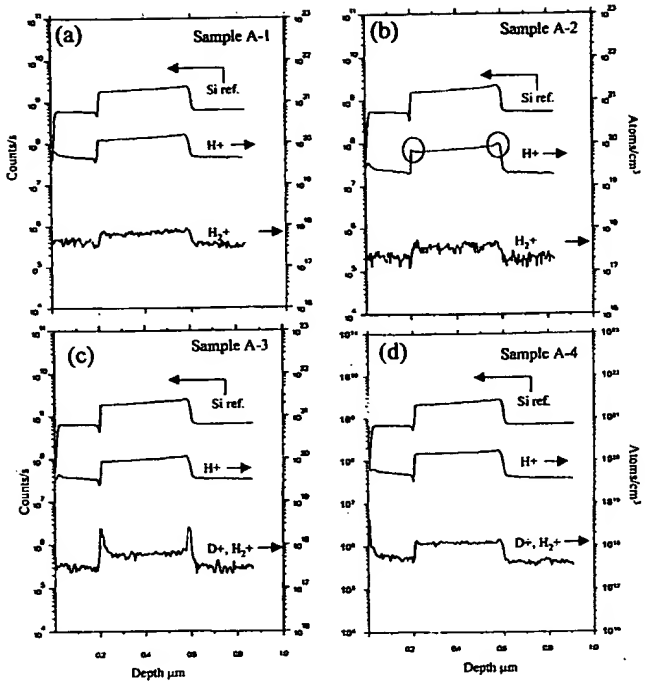


FIG. 4. SIMS depth profiles for (a) as-received (sample A-1), and after sintering in (b) $N_2$:$H_2$ at 625 °C for 1 h (sample A-2), (c) $N_2$:$D_2$ at 625 °C for 1 h (sample A-3), and (d) $N_2$:$D_2$ at 435 °C for 1 h (sample A-4).

plant (due to the large angle scattering of implanted $H^+$ near the end of range) distribution feature can be readily seen to span the interface with the abrupt ion yield enhancement described above [Fig. 2(a)] superposed upon this feature.

Examination of the same sample under positive secondary ion acquisition conditions results in the profile shown in Fig. 3(b). The corresponding implant distribution is observed, but is significantly broadened *within* the $SiO_2$ layer. Moreover, an *enhanced* concentration of H is observed in the shallow and deep interface region.

Similar results are observed for a $H^+$ implant tailored (30 keV, total dose $=3\times10^{15}/cm^2$) to exhibit a depth distribution peak at the center of the $SiO_2$ layer [Figs. 3(c) and 3(d)]. Again, the positive secondary ion profile [Fig. 3(d)] indicates that the original implant distribution within the $SiO_2$ layer is substantially perturbed in the positive ion acquisition mode, and a similar enhancement in H concentration is apparent at the Si/$SiO_2$ interface regions.

These results suggest that care must be taken in the interpretation of the SIMS profiles. We believe that the observed redistribution of the implanted hydrogen in the positive secondary ion acquisition mode results from the required SIMS experimental conditions. Apparently, the combined effect of $O_2^+$ ion beam bombardment and the charging compensation electron flux results in facile H transport/redistribution within the $SiO_2$ layer and a pileup of H at both interface regions (whereupon the abundant dangling bonds are passivated). We note, however, that the portion of the implant distribution within the Si region does *not* appear to be substantially perturbed from the acquisition process.

As we are primarily interested in the interfacial reactions in this study, we resort to acquiring profiles in the positive secondary ion mode to avoid an ion yield enhancement "artifact" in the near interface region. With the awareness of the potential spatial perturbation associated with the positive secondary ion profile acquisition conditions, we cannot uniquely establish the initial, unperturbed H or D concentration throughout the $SiO_2$ layer in the SOI structure from our SIMS measurements. Therefore, we limit our study mainly to the changes in the interfacial reaction behavior.

## B. Incorporation of H/D

Figure 4 shows the SIMS analysis of samples A-1, A-2, A-3, and A-4. Annealing the as-received sample [Fig. 4(a)] in forming gas at 625 °C for 1 h results in the profile shown in Fig. 4(b). The temperature employed for this anneal is consistent with a saturated (mobile) proton density in similar SOI material[9] and the observation of efficient hydrogen transport in undoped poly Si.[14] The relatively high $H^+$ background interferes with the observation of annealing-induced H accumulation at the interfaces, although a careful examination suggests that an increase in H concentration at the interfaces is discernible (circled areas) in comparing Figs. 4(b) and 4(a). This increase is consistent with the passivation of dangling bonds at the Si/$SiO_2$ interface produced from the reaction of $H_2$ during the annealing process.[4]

In order to facilitate the direct observation of the interface passivation by the SIMS technique, the hydrogen in the forming gas mixture was replaced with the isotope deuterium. A fresh as-received sample was annealed in $N_2$:$D_2$ at 625 °C for 1 h (sample A-3) and the SIMS results are shown

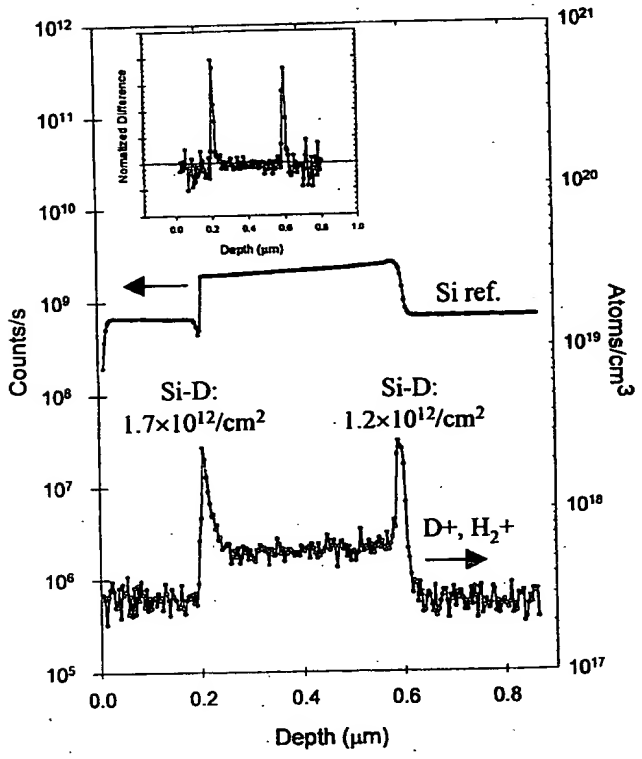


FIG. 5. Quantification of D concentration from sample A-3 at the $SiO_2$/Si interface.

in Fig. 4(c). It is evident that the D anneal at 625 °C has resulted in the incorporation of D at both Si/$SiO_2$ interfaces.

Comparison of the as-received and the $D_2$ annealed profiles also suggests that there is no increase in detectable D species against the $m/e=2$ background throughout the oxide layer. Previous SIMS work by others has also indicated that the D incorporation is largely confined to the interfacial region.[21] However, given the complications observed here associated with the possible redistribution of H or D in the positive secondary ion mode, one may well speculate that such species could be distributed uniformly throughout the oxide layer and piled up at the available sites in the interfacial region.[13,22]

We also note by comparing the data in Fig. 4(c) with Fig. 4(d) for sample A-4 that sintering at 625 °C is apparently more effective at D incorporation at the interfaces than sintering at 435 °C. This is consistent with our previous studies which have observed a diffusion barrier for deuterium in lightly doped ($\sim 10^{15}/cm^3$) polycrystalline Si in that the *c*-Si overlayer in the SOI structure here is similarly doped.[14]

More accurate quantification of the amount of D observed at the interface upon sintering at 625 °C was achieved by multiplexing only on $m/e=28$ ($Si^+$) and $m/e=2$ ($H_2^+$, $D^+$) ions, as shown in Fig. 5. By integrating the D concentration over the interface region (and assuming ion yield change can be accounted for by normalizing to Si signal), we determine that the areal density of atomic D at *each* interface is $\sim 1-2\times 10^{12}/cm^2$ which is somewhat larger ($\sim 2\times$) than the expected interface state density for a thoroughly (e.g., vacuum annealed) un-passivated SOI structure.[10,19,20] The inset of Fig. 5 presents the result of normalizing both sample A-1 and sample A-3 profiles to the respective Si reference signals as a point-by-point ratio. The difference between these profiles (i.e., $N_2$:$D_2$ sintered–as-received) is then plotted as shown. This procedure illustrates that the observed interface signal is indeed statistically significant, and also suggests that within the $SiO_2$ layer there appears to have $[D]<4\times 10^{12}/cm^2$, assuming the limit of D detection is $1\times 10^{17}/cm^3$. We note, however, that the normalization procedure provides only a rough estimate of the D areal density at the interfaces and the associated errors could account for a significant portion of the 2× factor observed.

### C. Low temperature thermal stability

Annealing a portion of sample A-3 at 125 °C for 3 h (sample "A-5") and again at 185 °C (sample "A-6") under flowing $N_2$ resulted in no detectable change in the D concentration in the interfacial region, as shown in Fig. 6. This is consistent with the expected thermal activation energy of Si–D bond scission and would indicate that the passivated interface is thermally stable under such conditions. One expects nearly identical thermal dissociation behavior for the Si–H bond as it is essentially chemically indistinguishable under thermal activation and the associated isotope effects are small.[23,24]

### D. Isotopic exchange

Re-annealing a portion of sample A-3 under $N_2$:$H_2$ (sample "A-7") at 625 °C resulted in the reduction of the previously observed D pileup in the interface regions, as shown in Fig. 7. This is consistent with replacement of D by H at both interfaces. It is noted that the deeper interface has apparently undergone a *less efficient* isotopic exchange (to form Si–H) as some $D^+$ species (likely bonded as Si–D) remain detectable after the exchange experiment.

To examine this issue further, a second series of samples were prepared and submitted to this isotopic exchange process. Samples, labeled B-1 through B-4 in Table I, were submitted to $N_2$:$H_2$ and $N_2$:$D_2$ sinters in an alternating sequence at 625 °C for 1 h. SIMS depth profiles of these samples are shown in Fig. 8.

Careful examination of the SIMS profile from sample B-1 [initial sinter in $N_2$:$H_2$–Fig. 8(a)] indicates the enhanced $H^+$ concentration at the Si/$SiO_2$ interfaces, again likely present as Si–H bonds, as was seen in Fig. 4(b). Exposure of sample B-1 to the $N_2$:$D_2$ ambient (sample "B-2") results in the profile shown in Fig. 8(b). It is now noted that a *more efficient* isotopic exchange (to form Si–D) is observed at the deeper interface where an enhanced D concentration is observed (compare to Fig. 7, sample A-7). Moreover, an inequivalent concentration of D is evident at each interface, in contrast to that observed after the $D_2$ anneal of an as-received sample [see Fig. 4(c), for example].

Re-sintering sample B-2 in $N_2$:$H_2$ (sample "B-3") results again in isotopic exchange at both interfaces [Fig. 8(c)], with the deeper interface exhibiting a less efficient exchange.

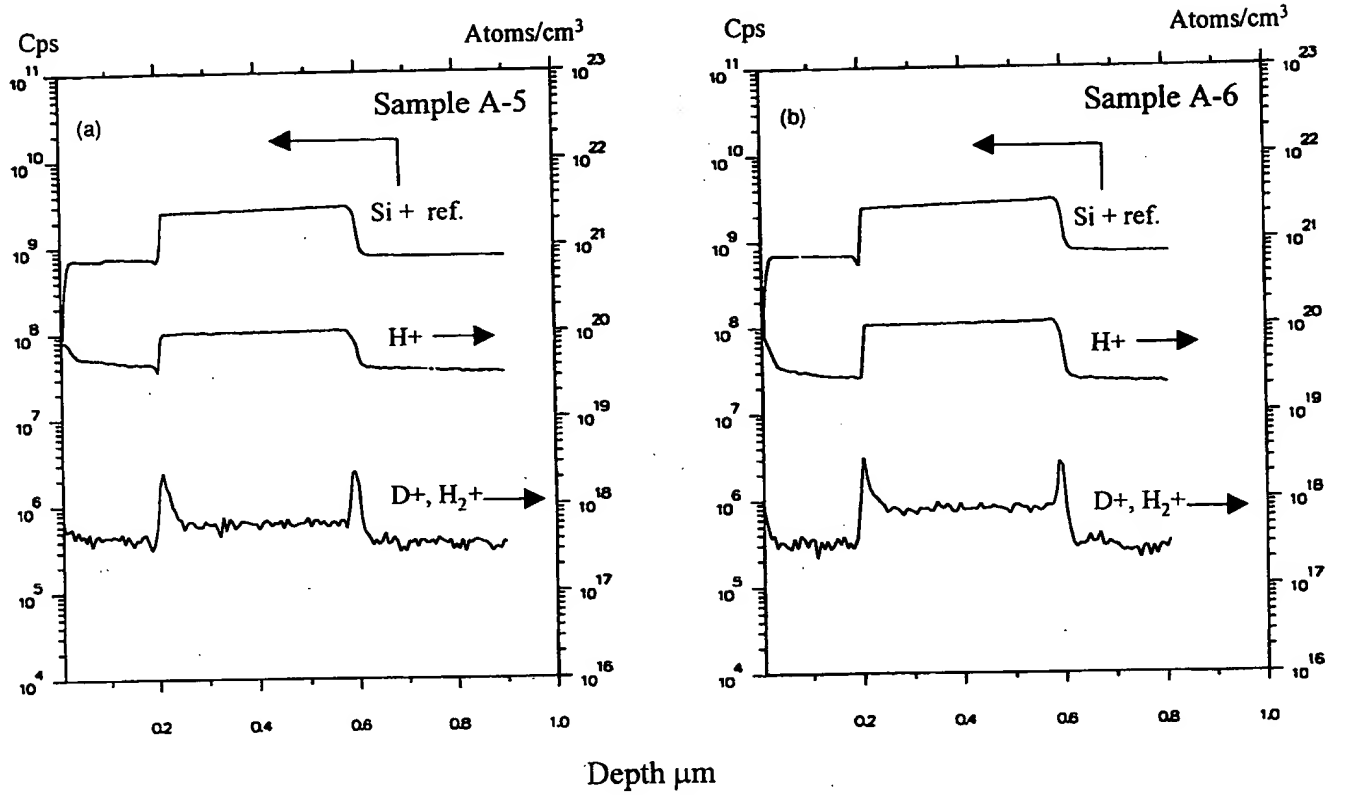


FIG. 6. SIMS profiles for $N_2$:$D_2$ annealed samples following lower temperature anneals in $N_2$ for 3 h: (a) 125 °C (sample A-5) and (b) 185 °C (sample A-6).

Careful examination of the $H^+$ profile again shows a discernible enhancement in the H concentration at the interface.

Finally, re-sintering sample B-3 in $N_2$:$D_2$ results in the profiles shown in Fig. 8(d) (sample "B-4"). Again, an enhanced D concentration is observed at both interfaces due to isotopic exchange, with more efficient exchange at the deeper interface.

To verify the reversibility of the observed exchange process, additional sintering experiments were also performed with the deuterium ($N_2$:$D_2$) exposure *first* in the sequence, followed by hydrogen ($N_2$:$H_2$), then deuterium ($N_2$:$D_2$), and finally $N_2$:$H_2$ again. Depth profiles of this sample sequence (shown in Fig. 9) indicate that complete exchange occurs at the shallow interface and again an enhanced concentration of D at the deeper interface was observed after sintering.

### E. Effect of cooling rate

We have performed preliminary measurements of the sensitivity of the quench rate to the H population as measured by SIMS at the interface. Figure 9(a) shows the positive secondary ion depth profile after rapid cooling (quenching) from 625 °C where an enhanced, roughly equal D concentration in each interface region is evident. Slowly cooling the sample in $N_2$:$D_2$ (sample "D-1") results in the SIMS profile shown in Fig. 10. This was done by annealing the sample at 625 °C in $N_2$:$D_2$ for 60 min, followed by slowly cooling to ~300 °C in a $N_2$:$D_2$ ambient, followed by further slow cooling to room temperature in a $N_2$ ambient in the same furnace.

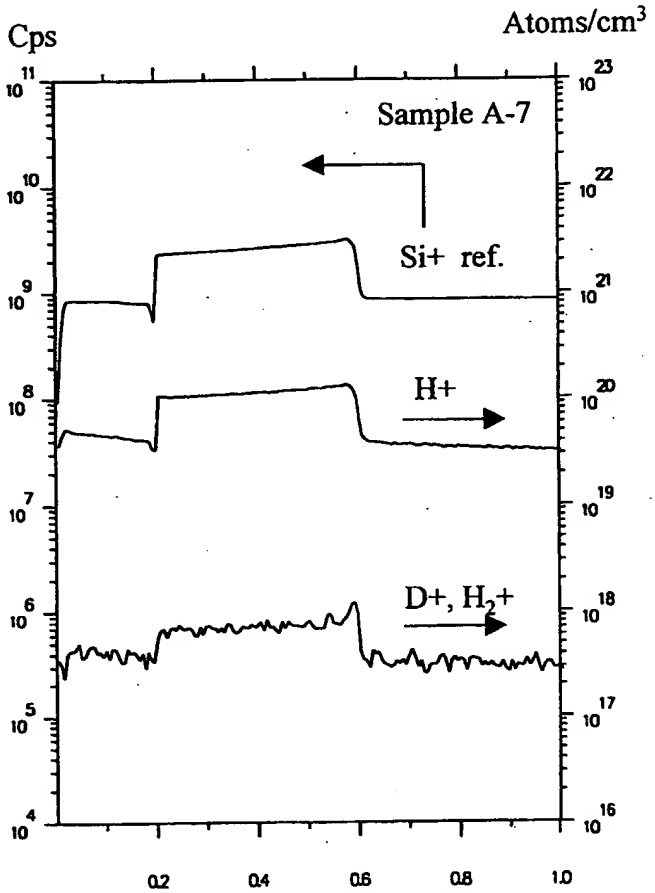


FIG. 7. SIMS profile demonstrating isotopic exchange of interface D with H (sample A-7).

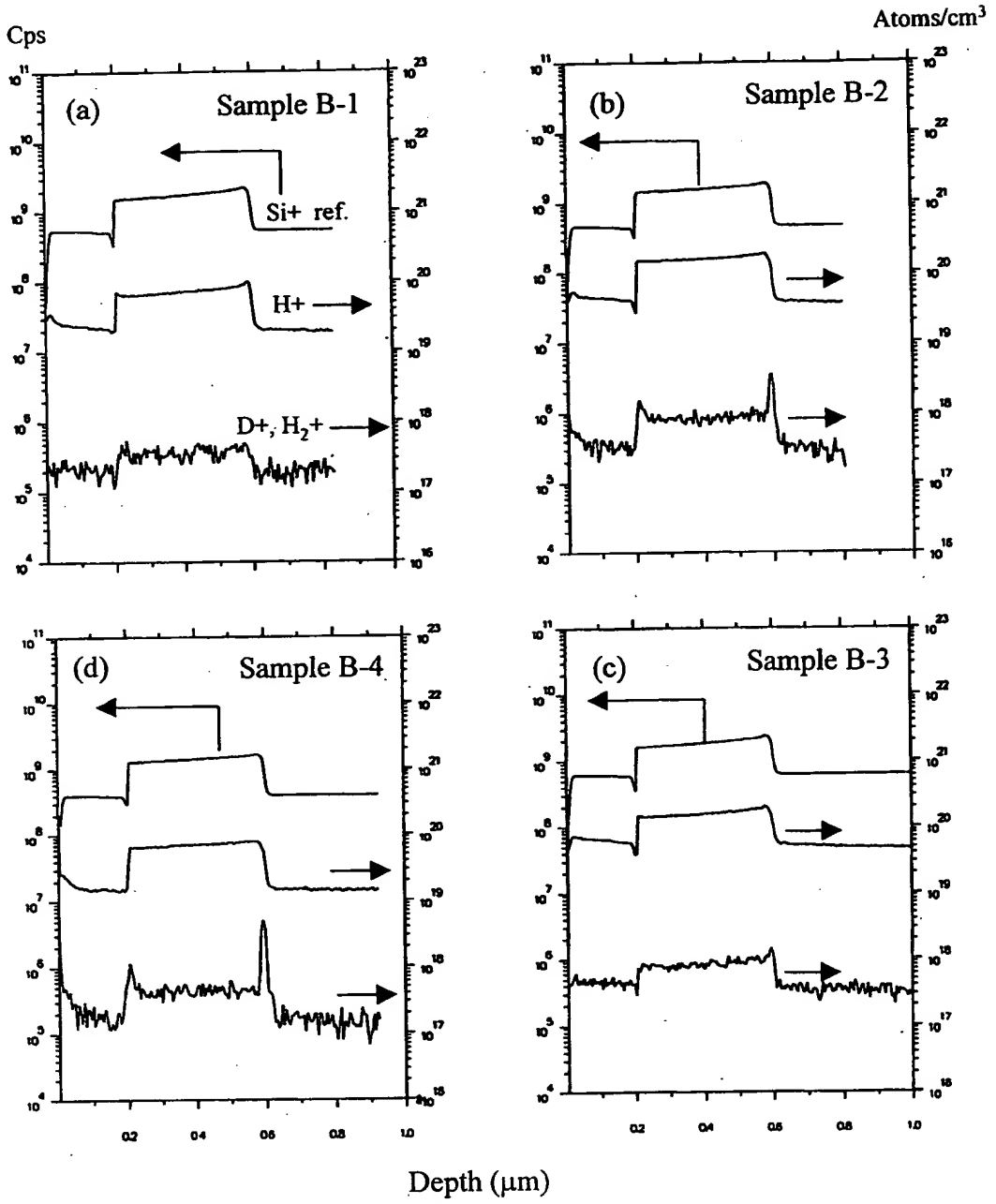


FIG. 8. SIMS depth profiles from isotopic exchange experiments (625 °C, 1 h). (a) Sample B-1, (b) sample B-2, (c) sample B-3, and (d) sample B-4.

An asymmetry in the D concentration is observed: a somewhat lower population is observed at the shallow interface in comparison to that observed at the deep interface.

## IV. DISCUSSION

To interpret these results, some familiarity with the Unibond® growth process is useful.[16] Briefly, a high quality ~400 nm thermal oxide is grown on a (first) Si substrate, and is subsequently implanted with hydrogen ions at a depth below the $Si/SiO_2$ interface. This results in the formation of microcavities with a volume on the order of the implant straggle. A (second) "handle" Si wafer (with a native oxide) and the first implanted/oxide wafer are then cleaned with a modified RCA process in order to render both oxide surfaces with a hydroxyl (hydrophilic) termination. These hydrophilic wafer surfaces are then placed together and initial bonding is facile through the hydrogen bonding of the surfaces. A first thermal treatment is performed at 400–600 °C in order to permit a crack formation among the H-filled microcavities in the first wafer and this results in the spalling of the first wafer substrate. Thus, a buried oxide with a single crystal Si overlayer ($p$ type, $5\times10^{15}/cm^3$) is left on the handle wafer. (The first wafer can then be polished and used again for subsequent SOI growth.) A final high temperature anneal (~1050–1350 °C) is performed on this structure to stabilize the bonding interface. The rough surfaces left over from the spalling step are then polished resulting in the removal of a few hundred angstroms of the Si overlayer. Interface state

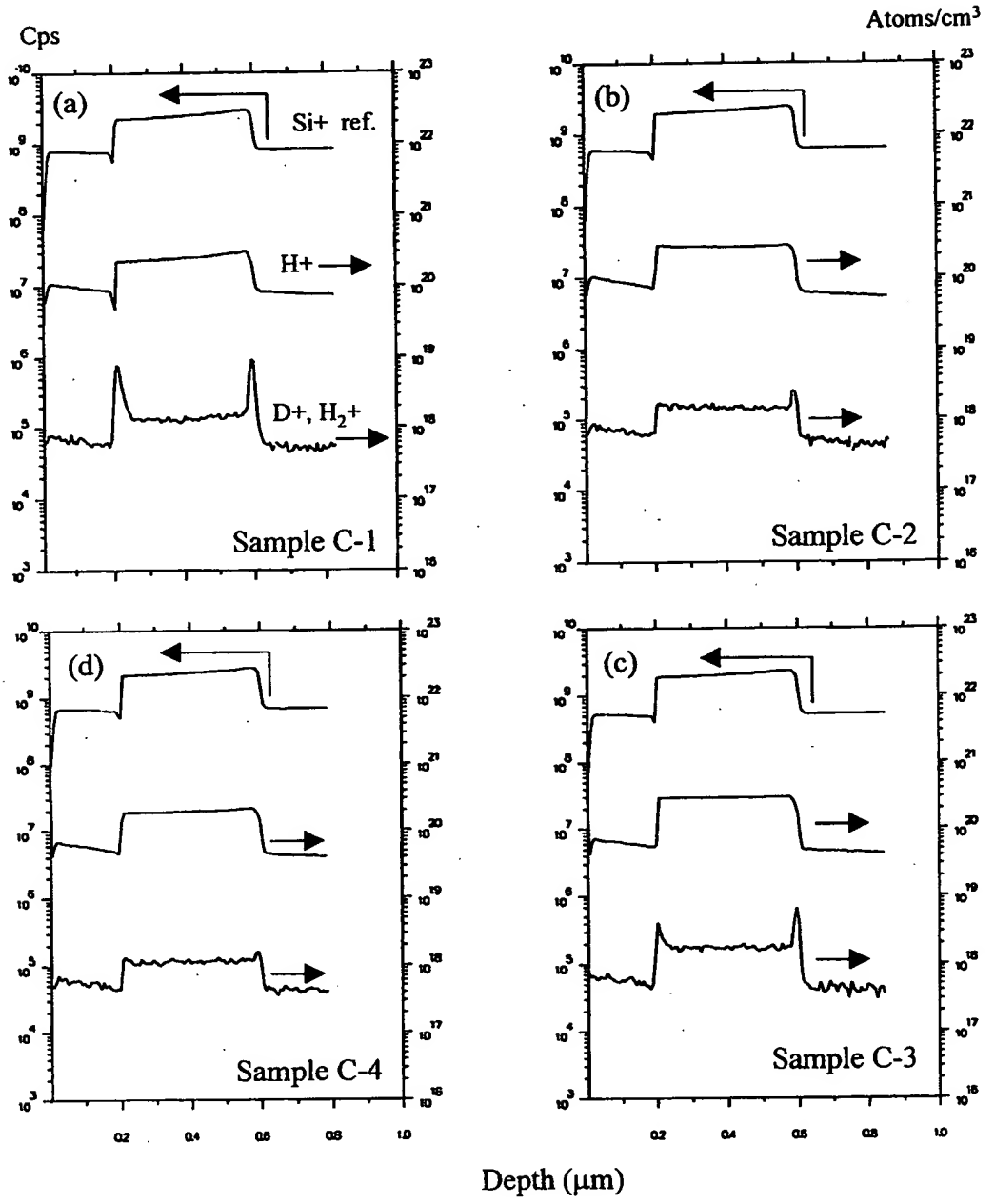


FIG. 9. SIMS depth profiles from isotopic exchange experiments (625 °C, 1 h). (a) Sample C-1, (b) sample C-2, (c) sample C-3, and (d) sample C-4.

densities of less than $10^{11}$/eV/cm$^2$ and a mobility of >650 cm$^2$/V/s have been reported for psuedo-metal-oxide-semiconductor field effect transistor (MOSFET) devices incorporating Unibond®.[16]

### A. Exchange at the $SiO_2$/Si interfaces

The observation of a preference for Si–D formation at the deeper interface (originally associated with the handle wafer surface) is intriguing. An obvious source of the difference in isotopic exchange behavior between the shallow and deep interfaces could be the Unibond® fabrication process itself. The shallow interface is resultant from a thermal oxidation process and such interfaces have been well characterized to exhibit Si dangling bonds which are largely passivated upon $H_2$ (or $D_2$) sintering at $T>400$ °C. The deeper Si/$SiO_2$ interface, however, was formed from the coupling of a hydrophilic native oxide to the (hydrophilic) thermal oxide surface followed by an anneal for crack formation. Of course the entire structure, and therefore both interfaces, is eventually submitted to an extremely high temperature, inert gas anneal.

Evidence for O vacancy production (strained Si–Si bond formation) from such processes in analogous oxygen-implanted separation by implantation of oxygen (SIMOX) structures has been reported.[5,25] It is noted that the associated interfaces in such SIMOX structures are expected to be initially equivalent, in contrast to the structure produced by the Unibond® process examined here. As the initial interfaces are inequivalent in Unibond® material, we speculate that the deeper interface is somehow rendered more sensitive to oxygen vacancy defect generation in comparison to the shallow

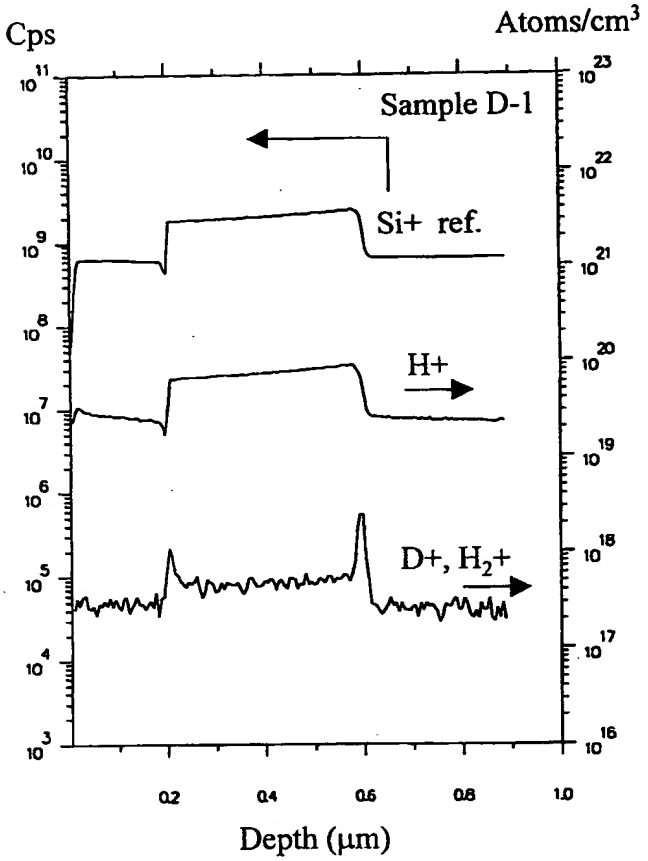


FIG. 10. SIMS depth profile after slow cool from 625 °C $N_2$:$D_2$ for 60 min (sample D-1).

interface, in spite of the subsequent, extensive thermal treatments. We further speculate that this fabrication process may result in the formation of a defect (of unknown bonding structure) that exhibits a kinetic isotope effect leading to the propensity to preserve the Si–D bond. It is also noted that such defects might also lead to a substantial difference in molecular $D_2$ scission efficiency relative to $H_2$ scission thus resulting in the preference for Si–D stability.

As the shallow interface is expected to exhibit the properties of conventional thermal oxides due to the fabrication sequence, a reversible isotopic exchange mechanism is anticipated at $T>600\,°C$ due to the essentially indistinguishable thermal activation of Si–H and Si–D bond scission. Our observations are consistent with such an interface model.

We note, however, that recent theoretical work has suggested that, for noninteracting adatoms on surfaces, the propensity for Si–D formation over Si–H formation on Si surfaces can be explained upon the consideration of the equilibrium lattice dynamics.[26] Our results indicate that interface structure is also apparently an important consideration in regard to efficient isotope switching under the equilibrium sintering conditions employed here.

### B. Relevance to mobile protons in buried $SiO_2$

We first emphasize that the SIMS results reported here have relevance to the mobile proton phenomenon reported in the literature.[6–9] To the best of our knowledge, the first report of an observation of H redistribution upon avalanche injection detected by SIMS and *CV* measurements was by Gale *et al.*[13] In that work, a linear increase in interfacial H concentration with electron injection current was presented and attributed to the formation of a mobile proton population.

The observation of H *implant* redistribution for the positive versus negative secondary ion profiles reported here clearly indicates that a population of the available H can apparently be forced to migrate throughout the buried $SiO_2$ layer. As we have not intentionally attempted to enable H redistribution as in the work by Gale *et al.*[13] or Vanheusden *et al.*,[19] we suspect that the SIMS ion acquisition conditions employed here, viz., a simultaneous energetic e-beam flux for charging compensation and incident $O_2^+$ ion beam, results in the formation of a mobile hydrogen population until bound at a suitable site.

We note that a very recent report indicates that implanted $H^+$ into an identical Unibond® SOI structure to that employed here does *not* result in a significant mobile proton density, as detected by pseudo-MOSFET $I-V$ hysterisis measurements, even upon annealing the implanted sample in Ar at 600 °C.[19] A $I-V$ hysterisis characteristic, and therefore a detectable mobile proton population, is only apparent upon further annealing in forming gas as previously reported.[6–9,19] We suspect that the apparent discrepancy in the results seen here with this recent report is most likely due to the extreme conditions employed to obtain our SIMS data. For example, the induced fields from the electron beam charge compensation may be large due to charge imbalance during the (dynamic) depth profiling. Of course, the substantial perturbation of the $SiO_2$ layer structure from the implant process also results in extensive defect formation which may well affect the efficiency of mobile proton formation or subsequent transport in pseudo-MOSFET measurements.[19]

A significant portion of the enhanced concentration of H(D) we observe by SIMS at the interface upon $D_2$ *sintering* most probably coincides with *strongly bound* Si–H(D) species. We noted above that the measured D concentration at *each* interface ($\sim 1-2\times10^{12}/cm^2$) is somewhat larger than reported interface state densities for a thoroughly unpassivated SOI structure.[10,19,20] Moreover, the anticipated mobile proton density from the thermal treatments employed here are expected to be $\sim 3\times10^{12}/cm^2$ from the literature.[9] As we expect the SIMS measurement to detect virtually all of the D present at such densities, this suggests that some fraction of the species we detect at the interfacial region could therefore consist of mobile protons or their precursors.[27] Further clarification of the identity of the detected D by SIMS is problematic given the complications discussed above.

The high temperature anneal required in the SOI fabrication process is thought to introduce oxygen vacancies in the $SiO_2$ layer near the interface region, resulting in the formation of strained Si–Si bonds.[5,20] However, a significant effect in enhancing the density of mobile protons within buried

$SiO_2$ layers involves the rapid quench of the SOI structure from the hydrogen ambient ($T>600\,°C$) to room temperature.[9]

A model which seems consistent with the results reported here and these previous observations would entail a largely kinetic process whereby transport of molecular $H_2$ through the structure is followed by molecular bond scission at a suitable site, such as interface dangling bonds.[28,29] The atomic H liberated is then available for interface passivation and, upon saturation of the accessible dangling bond population, is finally followed by stable (mobile) $H^+$ formation.

The efficient transport of hydrogen to the $Si/SiO_2$ interface region is critical to the production of mobile proton species. Previous work[14] using SIMS to characterize the transport of deuterium to the $Si/SiO_2$ interface suggests that transport is not efficient in polycrystalline Si with low dopant concentration ($\sim 10^{15}/cm^3$, similar to the *c*-Si material in the Unibond® material employed here) for temperatures $<450\,°C$ [see Fig. 4(d)]. However, at $T=625\,°C$, efficient transport through polycrystalline Si is observed. Recent work on hydrogen transport in *c*-Si indicates that an interstitial diffusion mechanism is observed at the temperatures employed here, and that tunneling transport occurs at very low ($<200$ K) temperatures.[30] Importantly, lateral diffusion through the $SiO_2$ film is also expected to be an important mechanism at these temperatures as well.[9,12] We therefore anticipate that at $T=625\,°C$, the hydrogen transport is facile through either lateral ($SiO_2$) pathways or vertical (Si) pathways. At lower temperatures, lateral diffusion through the $SiO_2$ layer should dominate the process assuming that access to the $SiO_2$ layer is uninhibited.

Thus, as a result of efficient $H_2$ transport and bond scission, the interface dangling bonds are initially passivated resulting in the saturation of strongly bound Si–H. The presence of the oxygen vacancy sites at the buried $Si/SiO_2$ interface region, also responsible for $H_2$ scission and thus the production of H (and $H^+$ at $T>450\,°C$), is expected to provide mobile protons within the $SiO_2$ layer.[7] The presence of Si–H (and possibly $H_2$) at the interface may then serve as a barrier to out-diffusion of the mobile protons thus "trapping" them within the $SiO_2$ layer.[31]

The significant effect of the cooling rate on mobile proton density has been reported.[9] The SIMS results reported here indicate that an asymmetry in the strongly bound Si–H(D) population at the interface is observed as a result of such cooling. A possible explanation of this effect is that the slow cool process from 625 °C results in a higher probability for bound Si–H to dissociate (because the thermal activation occurs over a relatively longer time period) resulting in a relative loss of $P_b$ center passivation over that obtained from a rapid quench process. For example, efficient Si–H bond scission has been observed from temperature dependent spin resonance studies for $T\geqslant 460\,°C$.[29] One may then speculate that any subsequent mobile protons produced would be scavenged by such sites or diffuse away to recombine to form $H_2$ thus resulting in a substantial deficiency in the mobile proton population. The enhanced concentration of hydrogenic species at the deeper interface may indicate the presence of defects which have a propensity for Si–H(D) bond formation, as discussed above.

## V. CONCLUSIONS

It is evident that following sintering in $N_2$:$D_2$ forming gas, Si–D enrichment is taking place in the $SiO_2$/Si interface region. No detectable D is observed throughout the buried $SiO_2$ layer. This provides direct evidence that $D_2$ diffusion through the SOI structure can readily take place under the sintering conditions (625 °C for 1 h), allowing deuterium to reach the $Si/SiO_2$ interface region. We also note that a significant barrier to deuterium diffusion is observed at a lower sintering temperature (435 °C for 1 h) which is consistent with our previously reported results on CMOS structures.

We estimate that the Si–D areal density values are observed to be $\sim 1-2\times 10^{12}/cm^2$ for both interfaces and, in view of the errors associated with SIMS quantification across the interface, are in good agreement with commonly accepted interfacial Si-dangling bond density ($D_H \sim 1\times 10^{12}\,cm^{-2}$).

Annealing a $D_2$ sintered sample at $T\leqslant 185\,°C$ for several hours results in no detectable change in the D concentration at the interface and is consistent with the expected thermal stability of the Si–D bond. Identical behavior is expected for Si–H bonds.

Isotopic exchange of D with H is readily accomplished at 625 °C and is also consistent with similar Si–H and Si–D bond dissociation energies. However, sequential isotopic exchange experiments indicate that the Unibond® SOI interface structures *are distinguishable*. The deeper interface appears to undergo an isotopic exchange process which prefers Si–D bond formation over that of Si–H bond formation. In contrast, the shallow $Si/SiO_2$ interface does not appear to exhibit a propensity for Si–D formation as isotopic exchange appears to be complete. The isotopic exchange process appears to be largely reversible. Samples initially submitted to a $N_2$:$H_2$ treatment exhibit similar D interface concentrations upon subsequent deuteration compared to samples initially submitted to a $N_2$:$D_2$ treatment that undergo subsequent hydrogenation.

Finally, the effect of cooling rate on the SIMS profiles indicates that a quench process after sintering results in a symmetric population of strongly bound Si–H(D) in the $Si/SiO_2$ interface region, whereas a slow cooling rate does not.

## ACKNOWLEDGMENTS

The authors wish to thank Norman James for the $D_2$ sintering of samples. Shaoping Tang, Dick Stolz, and David Denker are gratefully acknowledged for useful discussions. The authors also acknowledge their collaboration with Bill Warren of Sandia National Laboratories, Karel Vanheusden of the Air Force Research Labs, and Professor Q.-M. Zhang at the University of Texas at Arlington.

[1]A. Stesmans and K. Vanheusden, Proceedings of the IEEE International SOI Conference, 1992, p. 102.
[2]K. Vanheusden and A. Stesmans, Appl. Phys. Lett. **64**, 2575 (1994).
[3]K. Vanheusden, A. Stesmans, and V. V. Afanasev J. Appl. Phys. **77**, 2419 (1995); J. Non-Cryst. Solids **187**, 253 (1995).
[4]K. Vanheusden, W. L. Warren, J. R. Schwank, D. M. Fleetwood, M. R. Shaneyfelt, P. S. Winokur, and R. A. B. Devine, Appl. Phys. Lett. **68**, 2117 (1996).
[5]W. L. Warren, K. Vanheusden, J. R. Schwank, D. M. Fleetwood, M. R. Shaneyfelt, P. S. Winokur, and R. A. B. Devine, Appl. Phys. Lett. **68**, 2993 (1996); IEEE Trans. Nucl. Sci. NS-43, 2617 (1996), and references therein.
[6]K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur, and Z. J. Lemnious, Nature (London) **386**, 587 (1997); *Amorphous and Crystalline Insulating Thin Films—1996*, MRS Symposium Proceedings, edited by W. L. Warren, R. A. B. Devine, M. Matsumura, S. Cristoloveanu, Y. Homma, and J. Kanicki (Materials Research Society, Pittsburgh, 1997), Vol. 146, p. 187.
[7]W. L. Warren, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, B. L. Draper, P. S. Winokur, M. G. Knoll, K. Vanheusden, R. A. B. Devine, L. B. Archer, and R. M. Wallace, IEEE Trans. Nucl. Sci. NS-44, 1789 (1997).
[8]R. K. Lawrence, H. L. Hughes, R. E. Stahlbush, D. I. Ma, and M. E. Twigg, Proceedings of the IEEE International SOI Conference, 1997, p. 142.
[9]K. Vanheusden, W. L. Warren, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, B. L. Draper, P. S. Winokur, R. A. B. Devine, L. B. Archer, G. A. Brown, and R. M. Wallace, Appl. Phys. Lett. **73**, 674 (1998).
[10]K. Vanheusden, R. A. B. Devine, J. R. Schwank, D. M. Fleetwood, R. G. Polcawich, W. L. Warren, S. P. Karna, and R. D. Pugh, IEEE Trans. Nucl. Sci. NS-44, 2087 (1997).
[11]K. Vanheusden, S. P. Karna, R. D. Pugh, W. L. Warren, D. M. Fleetwood, R. A. B. Devine, and A. H. Edwards, Appl. Phys. Lett. **72**, 28 (1998).
[12]K. Vanheusden (private communication).
[13]R. Gale, F. J. Feigl, C. W. Magee, and D. R. Young, J. Appl. Phys. **54**, 6938 (1983).
[14]P. J. Chen and R. M. Wallace, Appl. Phys. Lett. **73**, 3441 (1998); Mater. Res. Soc. Symp. Proc. **513**, 325 (1998).
[15]SOITEC, 2 Centennial Drive Peabody, MA 01960. Phone (508) 531.2222.
[16]A. J. Auberton-Herve, T. Barge, F. Metral, M. Bruel, B. Aspar, C. Maleville, H. Moriceau, and T. Poumeyrol, in Ref. 6, p. 177.
[17]According to the spectra gases literature, the largest impurity concentration in pure $D_2$ is HD (~2900 ppm) followed by $H_2$, $N_2$, $D_2O$, $O_2$, $H_2O$ and hydrocarbons (all <20 ppm).
[18](a) H-implant in Si standard reference material, Charles Evans & Associates, Redwood City, CA; (b) H and D implanted in Si by IICO Corp., CA. The quoted hydrogen dose was cross calibrated with standard (a) and was found to be in agreement within 5%. The deuterium dose was used as quoted.
[19]K. Vanheusden, D. M. Fleetwood M. R. Shaneyfelt, B. L. Draper, and J. R. Schwank, IEEE Trans. Nucl. Sci. **45**, 2391 (1998).
[20]K. Vanheusden, W. L. Warren, W. M. Shedd, R. D. Pugh, D. M. Fleetwood, J. R. Schwank, and R. A. B. Devine, in Ref. 8, p. 64.
[21]H. Fukuda, T. Ueno, H. Kawarada, and I. Ohdomari, in *Control of Semiconductor Interfaces I*, edited by I. Ohdomari, M. Oshima, and A. Hiraki (Elsevier, New York, 1994), p. 211.
[22]S. M. Myers, J. Appl. Phys. **61**, 5428 (1987).
[23]H. Neergaard and J. T. Yates, Jr., Chem. Rev. **95**, 1589 (1995).
[24]C. G. Van de Walle Mater. Res. Soc. Symp. Proc. **513**, 55 (1998).
[25]R. A. B. Devine, D. Mathiot, W. L. Warren, and B. Aspar, J. Appl. Phys. **79**, 2302 (1996).
[26]I. P. Ipatova, O. P. Chikalova-Luzina, and K. Hess, J. Appl. Phys. **83**, 814 (1998).
[27]We note that measurements of interface trap densities $\sim 5\times10^{12}/cm^2$ have been reported for thermal oxides on Si. See Ref. 30 below.
[28](a) J. H. Stathis, in *The Physics and Chemistry of $SiO_2$ and the Si-$SiO_2$ Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993), p. 343; (b) A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter **9**, 3299 (1997), and references therein; (c) **10**, 19 (1998).
[29](a) K. L. Brower and S. M. Myers, Appl. Phys. Lett. **57**, 162 (1990); (b) E. Cartier, J. H. Stathis, and D. A. Buchanan, *ibid.* **63**, 1510 (1993); (c) A. Stesmans, *ibid.* **68**, 2076 (1996).
[30]S. Fabian, S. Kalbitzer, Ch. Klatt, M. Behar, and Ch. Langpape, Phys. Rev. B **58**, 16144 (1998).
[31]C. H. Seager and R. A. Anderson, Appl. Phys. Lett. **53**, 1181 (1988); C. H. Seager, R. A. Anderson, and J. K. G. Panitz, J. Mater. Res. **2**, 96 (1987).

# Reduction of hot electron degradation in metal oxide semiconductor transistors by deuterium processing

J. W. Lyding and K. Hess
*Department of Electrical and Computer Engineering and Beckman Institute, University of Illinois, Urbana, Illinois 61801*

I. C. Kizilyalli
*Lucent Technologies Bell Laboratories, Orlando, Florida 32819*



(Received 22 January 1996; accepted for publication 28 February 1996)

We report experimental results that replacing hydrogen with deuterium during the final wafer sintering process greatly reduces hot electron degradation effects in metal oxide semiconductor transistors due to a new giant isotope effect. Transistor lifetime improvements by factors of 10–50 are observed. A plausible physical theory suggests that the benefits of deuterium use may be general and also applicable to other areas of semiconductor device processing and fabrication. © *1996 American Institute of Physics.* [S0003-6951(96)00418-4]

The time-dependent degradation of metal oxide semiconductor (MOS) transistor performance resulting from hot (energetic) electron effects has been an area of considerable study over the past 25 years.[1] According to established theory, this aging process is thought to occur in part as the result of hot electrons stimulating the desorption of hydrogen from the $Si/SiO_2$ interface region. Hydrogen is introduced by necessity during several device processing steps as, for example, during the sintering of the wafers at elevated temperature in a hydrogen ambient.[2] While this process improves device function, it sets the stage for subsequent hot electron degradation. In this letter we demonstrate an alternative process in which the interface states are passivated by deuterium instead of hydrogen. Transistors that have been annealed with deuterium show a greatly reduced degradation due to hot electron effects.

The idea of using deuterium instead of hydrogen was in part inspired by experiments in which a scanning tunneling microscope (STM) was used to stimulate the desorption of hydrogen from Si(100)2×1:H surfaces under ultrahigh vacuum (UHV) conditions.[3] Following the suggestion of Avouris,[4] these experiments were extended to deuterated surfaces in order to explore more fully the surface science issues of this process. From these new experiments it was discovered that deuterium is much more difficult to remove under the conditions used to desorb hydrogen.[5] While there are clearly many differences between a free surface in UHV and a buried $Si/SiO_2$ interface, this result suggests the possibility for a sizable isotope effect if hydrogen is replaced by deuterium during the conventional wafer sintering step. To test for the advantages of using deuterium, uncapped complementary metal oxide semiconductor (CMOS) wafers fabricated at Bell Laboratories were subjected to the deuterium sintering process at Illinois and then returned to Bell Laboratories for electrical stress testing.

The wafers used for our tests contained *n*-channel metal oxide semiconductor (NMOS) transistor structures fabricated using the Bell Laboratories 0.5 $\mu$m 3.3 V CMOS technology.[6] However, the following three changes were made: (i) the gate oxide was reduced to $t_{ox} \sim 55$ Å, (ii) the doping in the *p*-well was increased, and (iii) the phosphorus doped lightly doped drain region was replaced by a shallow arsenic implanted (dose = $4\times10^{14}$ cm$^{-2}$ and energy = 30 keV) source-drain extension region. These process modifications enhance the peak value for the source-drain electric field near the drain edge of the gate, resulting in more channel hot electrons. The shallow source-drain extension ensures that these hot electrons are near the $Si/SiO_2$ interface, where they will cause significant interface damage. The interface damage, caused by these hot carriers, can easily be observed by monitoring the change in the NMOS transistor transconductance (i.e., $g_m = \Delta I_{DS}/\Delta V_{GS}|_{V_{DS}=0.1\ \mathrm{v}}$) or by the shift in the transistor threshold voltage $V_{TH}$.[7]

For this study, accelerated hot carrier dc stress experiments were performed on transistors with varying gate lengths (0.5–15 $\mu$m) at peak substrate current conditions. The applied (accelerated) stress source drain voltage was $V_{DS}=5$ V and the source gate voltage was $V_{GS}=2$ V. Stress experiments performed at lower voltages ($V_{DS}=3.8$ V and $V_{GS}=1.5$ V) and shorter gate lengths (0.3 and 0.4 $\mu$m) give results similar to the ones reported below. Prestress transistor measurements demonstrate that devices sintered in hydrogen or deuterium have identical electrical characteristics (e.g., transconductance, subthreshold slope, threshold voltage, saturation current, substrate current, etc.).

Figure 1 shows the $g_m$ degradation as a function of stress time for NMOS transistors with five gate lengths ranging from 0.5 to 0.7 $\mu$m. Figure 2 shows the threshold voltage increase as a function of stress time for the same devices. All of these transistors are from the same wafer and were processed identically except for the manner in which they were sintered. Wafers sintered in deuterium exhibit much more resilience to channel hot carrier stress. In our comparative study, we have electrically stressed 80 or so transistors, and have observed the same strong trend. These results have also been verified by performing the same electrical stress experiments on a second wafer from another lot. If we use 20% $g_m$ degradation as a lifetime criterion, transistors sintered in deuterium typically have lifetimes 10–50 times longer than those sintered in hydrogen. Likewise, we observe a factor of 10 improvement in lifetime if we take a shift of 200 mV in threshold voltage as the degradation criterion. In our opinion,

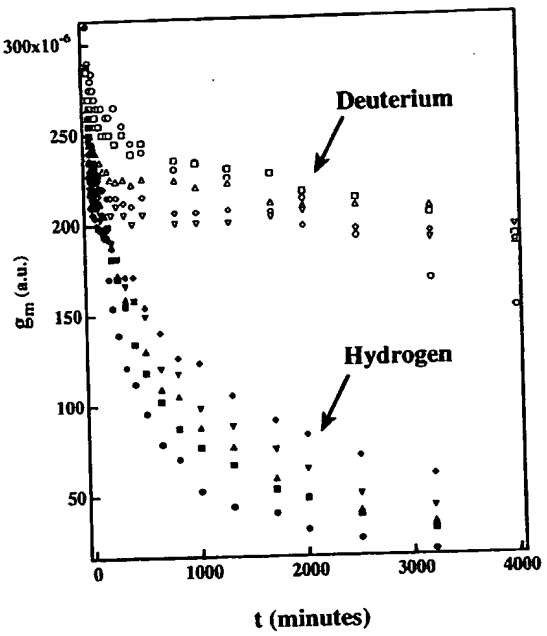


FIG. 1. Comparative time-dependent degradation of the transconductance $g_m$ for five NMOS transistors sintered in hydrogen (solid symbols) and deuterium (unshaded symbols). The sinter process was performed in a 90% $N_2$:10% $H_2(D_2)$ ambient at 400 °C for 1 h.

further improvements in lifetime will be achieved once the optimum sinter process conditions are determined.

The theoretical explanation of the reduced hot electron degradation due to deuterium treatment is probably analogous to the explanation for the STM experiments by Avouris *et al.*,[8] although this analogy should not be pushed too far. They showed that hydrogen absorbed on the silicon surface can be taken off by a STM tip up to 100 times easier than deuterium. An explanation of such a giant isotope effect can be found by assuming that the hot electrons cause a population of a silicon hydrogen (deuterium) antibonding state. This results in a force that accelerates the hydrogen away from the silicon surface leading to reduced wave function overlap with bulk silicon states. That acceleration is, of course, much diminished for the deuterium because of its bigger mass. For the same reason the kinetic energy gain necessary for dissociation is reached faster by hydrogen than by deuterium. In other words, the differences between hydrogen and deuterium arise from dynamic effects as they are important in chemical reactions. The static chemical bonding is evidently the same for both hydrogen and deuterium which is the reason for the identical transistor properties after hydrogen and deuterium treatment before hot electron dynamics and resultant damage. The difference of the hot electron degradation compared to the STM experiment lies mainly in the more complex chemistry of the $Si/SiO_2$ interface. The hydrogen (deuterium) passivating a silicon bond may (due to hot electron excitation) transfer to the $SiO_2$, passivating a more removed silicon bond or linking up with oxygen or even forming $H_2$ ($D_2$). All of these processes may be complicated by processes of interface reconstruction and defect chemistry. As a consequence, the energy needed to depassivate and remove hydrogen (deuterium) may be significantly different from the energies in the STM experiments. The close proximity of $SiO_2$ (instead of the more remote tip) will also permit different reactions of hydrogen (deuterium) than the "reaction" with the tip electrode.

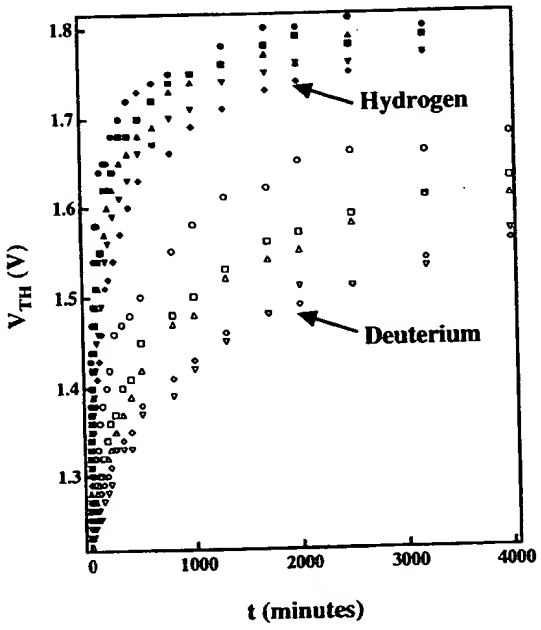


FIG. 2. Comparative time-dependent increase of the threshold voltage $V_{th}$ for NMOS transistors sintered in hydrogen (solid symbols) and deuterium (unshaded symbols). These are the same five devices shown in Fig. 1.

In addition to the effects discussed above there are also other effects that may explain the large improvement of hot electron degradation by use of deuterium. One effect is the well-known isotope effect that relates to the larger zero point energy of the Si–H oscillations as compared to Si–D. Another effect is the possibility of excited Si–H or Si–D oscillations. Hot electrons in devices may excite by multiple impact oscillations far above the thermal equilibrium and thus force dissociation. A multiple excitation mechanism is also thought to explain hydrogen desorption at lower electron energies in the STM experiments.[9] Again, deuterium would show less energy transfer because of its large mass. Another explanation is related to the possible mass dependence of tunneling of the nuclei that might be involved in the chemical process of dissociation. We do not want to speculate at this point on which effect is dominant. All of these effects favor deuterium as the more stable passivation. The generality of these effects suggests that deuterium instead of hydrogen may be beneficial in other processes, devices, and device materials. Note that the well-known effects based on differing mobilities of $H^+$ and $D^+$ (e.g., following gamma irradiation[10]) or as occurring in electrolysis will not explain the giant isotope effect that we observe. These effects do not involve hot electrons at the interface.

In conclusion, we have demonstrated that the replacement of hydrogen with deuterium during the final wafer sintering process results in substantially reduced susceptibility to hot electron degradation effects. The explanation of the effect is based on the increased difficulty to break the deuterium bond due to the additional neutron mass.

This work was supported by the Office of Naval Research University Research Initiative under Grant N00014-92-J-1519 and by the Beckman Institute for Advanced Science and Technology at the University of Illinois. The use of the AT&T Silicon Fabrication Facility in Orlando (OR-1) is gratefully acknowledged. We also acknowledge many valuable discussions with Dr. Jiri Jonas. (J. L.) would like to thank Dr. Ph. Avouris, G. Abeln, and Dr. T.-C. Shen for valuable discussions and collaboration in the STM experiments. (K. H.) would like to thank Dr. P. von Allmen, Dr. T. L. Brown, Dr. G. J. Iafrate, Dr. E. H. Poindexter, and Dr. P. Wolynes for valuable discussions. (I.C.K.) thanks Dr. K.-H. Lee, Dr. G. Higashi, and Dr. D. P. Chesire for their encouragement.

[1] S. Wolf, *Silicon Processing for the VLSI Era* (Lattice, Sunset Beach, 1995), Vol. 3, Chap. 9, pp. 559–674.
[2] A. B. Fowler, U.S. Patent No. 3,849,204 (1974).
[3] J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, Appl. Phys. Lett. **64**, 2010 (1995).
[4] Ph. Avouris and J. W. Lyding (private communication).
[5] G. C. Abeln, T.-C. Shen, Ph. Avouris, and J. W. Lyding (unpublished).
[6] I. C. Kizilyalli, M. J. Thoma, and F. L. Lytle, IEEE Trans. Semicond. Manuf. **8**, 440 (1995).
[7] J. M. Pimbley, M. Ghezzo, H. G. Parks, and D. M. Brown, in *Advanced CMOS Process Technology* (Academic, San Diego, 1989), Vol. 19.
[8] Ph. Avouris, R. E. Walkup, A. R. Rossi, H. C. Akpati, P. Nordlander, T.-C. Shen, G. C. Abeln, and J. W. Lyding, Surf. Sci. (to be published).
[9] T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, Ph. Avouris, and R. E. Walkup, Science **268**, 1590 (1995).
[10] N. S. Saks and R. W. Rendell, IEEE Trans. Nucl. Sci. **39**, 2220 (1992).

# Process Stability of Deuterium-Annealed MOSFET's

W. F. Clark, T. G. Ference, T. B. Hook, K. M. Watson, S. W. Mittl, and J. S. Burnham



*Abstract*— **Recent studies have shown improved NMOSFET transistor hot-carrier lifetime with the inclusion of a low-temperature post-metal anneal in a deuterium ambient. There have been few published results, however, on the optimization of the deuterium anneal or on the effect of further processing on deuterium-annealed samples. This paper reports the first results incorporating deuterium anneals at earlier steps in the process and the stability of the deuterium effect with further wafer processing. We also examine the effects of varying deuterium concentration in the anneal from 10 to 100% and annealing at different temperatures. Finally, the effect of combining a deuterium anneal with nitrided gate oxide is discussed.**

*Index Terms*—**Annealing, deuterium, hot carriers, MOSFET's, semiconductor processing.**

## I. Introduction

THE IMPACT of hot-carrier degradation on NMOSFET performance has been studied extensively since its discovery in the late 1970's [1]. The degradation has been correlated to the removal of hydrogen from the $Si/SiO_2$ interface due to collisions between heated carriers and the interface [2]. Recent studies indicate this degradation can be reduced when the typical hydrogen ambient post-metal anneals are done in a 10% deuterium/90% nitrogen ambient; in their model the authors attribute this improvement to the replacement of hydrogen at the interface by deuterium [3]. The deuterium is more difficult to remove under hot-carrier stress conditions. In this paper we confirm these initial results and examine this hot-carrier lifetime with higher-concentration deuterium anneals. We have annealed MOSFET's in a deuterium ambient earlier in the process sequence and have subjected samples annealed in deuterium to further processing to examine the extendibility of this technique to multilevel metal processes. We also examined the effect of combining a deuterium anneal with a nitrided gate oxide process.

## II. Experiments and Discussion

The device experiments were conducted on wafers processed with either a conventional 3.3 or 2.5 V CMOS technology [4], [5]. The samples were subjected to accelerated hot-carrier stressing. Device degradation was quantified by monitoring changes in saturated drain current and threshold voltage as a function of stress time. Oxide monitor wafers (12.5 nm thermal oxide on a p-type wafer) were also used to examine effects of anneal temperature and concentration. Secondary ion mass spectroscopy (SIMS) was conducted on all samples to verify the presence of deuterium in the films. The analysis was performed with a Cameca IMS-4f. Negatively charged hydrogen ($^1H^-$) or deuterium ($^2H^-$) secondary ions were detected after being desorbed by a 14.5 KeV $Cs^+$ primary ion beam. Interference from diatomic hydrogen ($^1H_2$) is minimal in this mode and deuterium naturally associated with hydrogen is low in the gate region. These interferences can be checked by comparing deuterium and hydrogen profiles in deuterated samples and hydrogen controls. Table I summarizes the sample anneal conditions.

NMOSFET's from group A were annealed after the first level of metal was defined. Samples from cell A1 were annealed in 10% deuterium/90% nitrogen at 400°C for 60 min. The control samples (A2) were annealed in 10% hydrogen/90% nitrogen. These NMOSFET's incorporated a silicon nitride layer between the silicon and the phosphosilicate glass layer (PSG). Device hot-carrier stressing was conducted; however, no improvement in transistor hot-carrier lifetime was observed. SIMS analysis confirmed low levels of deuterium in the gate oxide region; the silicon nitride layer acts as a barrier to deuterium [6].

To quantify the effect of anneal temperature on deuterium incorporation, oxide monitor wafers were studied. Wafers B1, B2 were annealed in 10% deuterium/90% nitrogen for 60 min. at 400 and 600 °C, respectively. Wafer B3 was annealed in 100% deuterium for 60 min. at 400 °C. SIMS analysis indicated the 600 °C anneal sample contained a deuterium dose of $1 \times 10^{13}$ atoms/cm$^3$. while both 400°C samples contained doses of $6 \times 10^{12}$ atoms/cm$^3$. Within the limits of the analysis, the sample profiles of the deuterium were similar. Although post-metal anneals are generally limited to 400 °C, the SIMS results suggest that premetal anneals would be more efficient at the higher temperatures.

NMOSFET's from groups C, D, and E were annealed following silicide formation and prior to any silicon nitride or PSG deposition. Group C samples (C1, C3, C6) were annealed in 10% deuterium/90% nitrogen while those in cell D1 were annealed in 100% deuterium. Typical hot-carrier stress results are illustrated in Fig. 1. The hot-carrier lifetime is improved by almost 30 times with the 10% deuterium-annealed NMOSFET's. Most of the control samples were annealed in a 10% hydrogen/90% nitrogen. The C2 cell was annealed in 5% hydrogen/95% nitrogen. Prior stress experiments varying hydrogen anneal concentrations from 5 to 100% showed equivalent hot-electron lifetime.

Fig. 2 depicts the lifetime to a 5 mV threshold voltage degradation for the various groups of samples. NMOSFET's

Manuscript received June 9, 1998; revised September 17, 1998.
The authors are with IBM Microelectronics Division, Essex Junction, VT 05452 USA (e-mail: a617853@us.ibm.com).
Publisher Item Identifier S 0741-3106(99)00420-6.

TABLE I
TECHNOLOGY TYPE AND ANNEAL CONDITIONS FOR EACH GROUP

| Group | Cell | Technology | Type of Anneal | Anneal Conditions | | | Sample Types | Results Summary |
|---|---|---|---|---|---|---|---|---|
| | | | | Ambient Gas | Temperature (°C) | Time (min) | | |
| A | 1 | 2.5V | Post 1st Metal | 10%$D_2$/90%$N_2$ | 400 | 60 | Un-nitrided Gate Oxide | No Improvement |
| | 2 | 2.5V | Post 1st Metal | 10%$H_2$/90%$N_2$ | 400 | 60 | Un-nitrided Gate Oxide | A1 Control |
| B | 1 | N/A | Post Oxide | 10%$D_2$/90%$N_2$ | 400 | 60 | Oxide Monitor Wafer | 6 x $10^{12}$ atoms/cm$^3$ |
| | 2 | N/A | Post Oxide | 10%$D_2$/90%$N_2$ | 600 | 60 | Oxide Monitor Wafer | 1 x $10^{13}$ atoms/cm$^3$ |
| | 3 | N/A | Post Oxide | 100%$D_2$ | 400 | 60 | Oxide Monitor Wafer | 6 x $10^{12}$ atoms/cm$^3$ |
| C | 1 | 2.5V | Post Silicide | 10%$D_2$/90%$N_2$ | 600 | 30 | Un-nitrided Gate Oxide | ~10X Lifetime (ΔVt) |
| | 2 | 2.5V | Post Silicide | 5%$H_2$/95%$N_2$ | 600 | 30 | Un-nitrided Gate Oxide | C1 Control |
| | 3 | 3.3V | Post Silicide | 10%$D_2$/90%$N_2$ | 600 | 30 | Nitrided Gate Oxide | ~10X Lifetime (ΔVt) |
| | 4 | 3.3V | Post Silicide | 10%$H_2$/90%$N_2$ | 600 | 30 | Nitrided Gate Oxide | C3, C5, D1 Control |
| | 5 | 3.3V | Post Silicide | 10%$H_2$/90%$N_2$ | 600 | 30 | Un-nitrided Gate Oxide | ~0.33X Lifetime (ΔVt) vs Nitrided Gate |
| | 6 | 2.5V | Post Silicide | 10%$D_2$/90%$N_2$ | 400 | 60 | Un-nitrided Gate Oxide | ~29X Lifetime (ΔIDsat) |
| | 7 | 2.5V | Post Silicide | 10%$H_2$/90%$N_2$ | 400 | 60 | Un-nitrided Gate Oxide | C6 Control |
| D | 1 | 3.3V | Post Silicide | 100%$D_2$ | 400 | 30 | Nitrided Gate Oxide | ~20X Lifetime (ΔVt) |
| E | 1 | 3.3V | Post Silicide | 10%$D_2$/90%$N_2$ | 400 | 60 | w/o Nitride Cap | E2 Control |
| | 2 | 3.3V | Post Silicide | 10%$D_2$/90%$N_2$ | 400 | 60 | w/ Nitride Cap | ~300X Reduction in D2 at Interface |

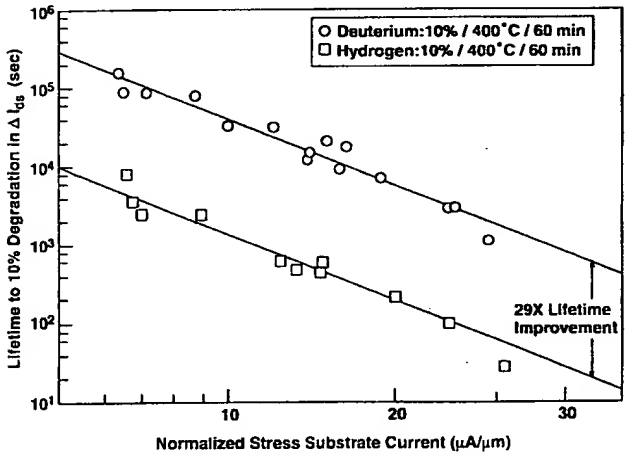


Fig. 1. Lifetime to a 10% degradation in reverse saturated drain current versus normalized substrate current for NMOSFET's annealed in a 10% deuterium ambient. The controls were annealed in a 10% hydrogen ambient. Stressing was conducted at maximum substrate current conditions for 24 h. The samples were annealed following silicide formation.

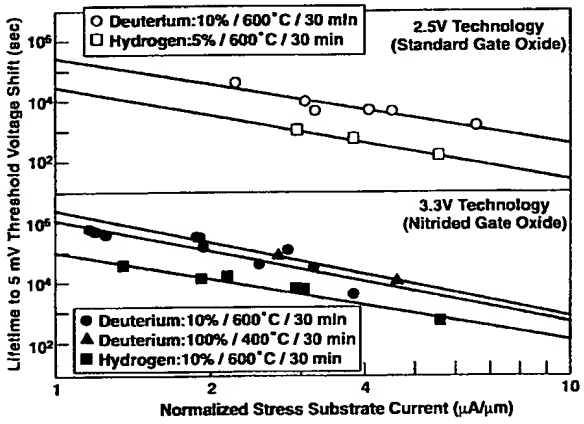


Fig. 2. Lifetime to 5 mV threshold voltage degradation versus substrate current for NMOSFET's annealed in 10% (C1—o, C3—•) and 100% (D1—▲) deuterium ambient. The 5 mV point was chosen due to the reduced degradation of the deuterium-annealed samples. The upper graph shows a comparison for the 2.5 V technology, incorporating a conventional gate-oxide, between 10% deuterium anneal (C1—o) and 5% hydrogen anneal (C2—□); a 10 times improvement in lifetime is observed. The lower graph shows a similar comparison for the 3.3 V technology, incorporating a nitrided gate-oxide process, between 10% deuterium anneal (C3—•), 100% deuterium anneal (D1—▲) and a 10% hydrogen anneal (C4—■). The 3 times improvement on the base (C4—■ versus C5) is attributed to the nitrided gate-oxide. The observed additional 10–20 times improvement in lifetime (C3—• and D1—▲ versus C4—■) is attributed to the deuterium annealing.

annealed in deuterium (C1) exhibit approximately 10 times lifetime improvement over those annealed in hydrogen (C2). This improvement is also observed on NMOSFET's with nitrided gate oxide (C3, D1) even though these transistors already exhibited a 3 times lifetime improvement over the un-nitrided oxide base (C5) [7]. Even with a lower temperature (400 °C) 100% deuterium anneal (D1), NMOSFET's exhibit approximately a 2 times improvement in lifetime over those annealed in a 10% deuterium anneal (C3) at 600 °C. We believe this apparent contradiction with the above SIMS results is due to uncertainties in extracting device lifetimes at post-silicide probe. Further work is required to fully quantify the concentration effect.

Several of these post-silicide annealed wafers (E1) were processed to first metal, including an intermediate silicon nitride and PSG deposition. At first-metal stress, no hot-carrier improvement was observed. As in group A, SIMS analysis of these samples confirmed low levels of deuterium in the gate oxide regions.

The SIMS analysis for the 10% deuterium anneal sample without subsequent processing (E1) is given in Fig. 3. The deuterium profile indicates low signal in the silicide with concentrations between $10^{18}$ and $10^{19}$ atoms/cm$^3$ in the polysilicon and gate oxide regions. The hydrogen profile in these regions essentially corresponds to the background level of the instrument. Samples from cell E1 were then subjected

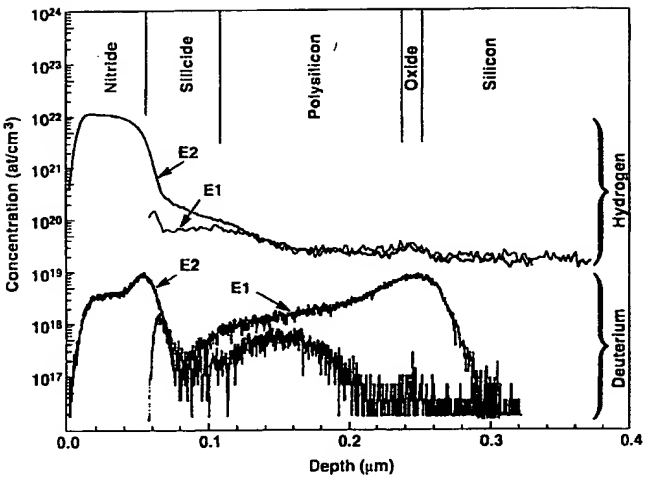


Fig. 3. SIMS profiles for samples annealed in 10% deuterium/90% nitrogen (E1) and one subsequently subjected to a silicon nitride deposition (E2) illustrating the migration of deuterium away from the oxide–silicon interface. The deposition was at 480 °C for 2 min. The position of the oxide layer was determined from profiles of $^{18}O$.

to a relatively low-temperature silicon–nitride deposition of 480 °C for 2 min. (E2); the SIMS analysis is included in Fig. 3 and shows that the deuterium present at post-silicide has migrated away from the $Si/SiO_2$ interface region. Significant deuterium levels are observed in the nitride film along with expected hydrogen levels. Comparison with the control suggests the majority of this deuterium is from the natural isotopic concentration of the hydrogen-rich film. However, the peak observed at the nitride/silicide interface is not present in the deuterium profile of the control sample. Evidently, this peak is related to the migration of deuterium from the $Si/SiO_2$ interface regions. Clearly, the nitride-cap processing is sufficient to modify the deuterium distribution in the films.

## III. Conclusion

We can improve NMOSFET hot-carrier lifetime at an earlier stage in the semiconductor process flow by annealing the transistor prior to silicon–nitride deposition. The increased device lifetime observed with nitrided gate oxides are further improved with deuterium anneals. The benefit of the deuterium anneal, however, is inherently unstable with further wafer processing. We also show an apparent improvement in hot-carrier lifetime with increased anneal temperature and deuterium concentration. These results provide evidence of significant challenges to the integration of deuterium anneal processing into mainstream semiconductor manufacturing.

## Acknowledgment

The authors would like to acknowledge those people from IBM's fabrication line in Essex Junction, VT who provided technical assistance. They also acknowledge the assistance of, and discussion with, J. W. Lyding and K. Hess of the Beckman Institute, University of Illinois, Urbana.

## References

[1] C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. W. Terrill, "Hot-electron-induced MOSFET degradation—Model, monitor, and improvement," *IEEE Trans. Electron Devices.*, vol. ED-32, pp. 375–385, 1985.

[2] D. A. Buchanan, A. D. Marwick, D. J. DiMaria, and L. Dori, "Hot electron induced hydrogen redistribution and defect generation in metal oxide semiconductor capacitors," *J. Appl. Phys.*, vol. 76, no. 6, pp. 3595–3608, 1994.

[3] J. W. Lyding, K. Hess, and I. I. Kizilyalli, "Reduction of hot electron degradation in MOS transistors by deuterium processing," *Appl. Phys. Lett.*, vol. 68, pp. 2526–2528, 1996.

[4] T. B. Hook, J. Piccirillo, K. Watson, and E. J. Nowak, "A CMOS technology for a sub-5ns 3.3V LVTTL 1-Mbit SRAM," in *Proc. 25th Eur. Solid State Device Research Conf.*, 1995, pp. 25–27.

[5] C. W. Koburger, W. F. Clark, J. W. Adkisson, E. Adler, P. E. Bakeman, A. S. Bergendahl, A. B. Botula, W. Chang, B. Davari, J. H. Givens, H. H. Hansen, S. J. Holmes, D. V. Horak, C. H. Lam, J. B. Lasky, S. E. Luce, R. W. Mann, G. L. Miles, J. S. Nakos, E. J. Nowak, G. Shahidi, Y. Taur, F. R. White, and M. R. Wordeman, "A half micron CMOS logic generation," *IBM J. Res. Develop.*, vol. 39, nos. 1/2, pp. 215–227, 1995.

[6] I. I. Kizilyalli, J. W. Lyding, and K. Hess, "Deuterium post-metal annealing of MOSFET's for improved hot carrier reliability," *IEEE Electron Device Lett.*, vol. 18, pp. 81–83, 1997.

[7] T. B. Hook, K. Watson, E. Lee, D. Martin, R. Ganesh, S. Kim, and A. Ray, "Correlation between fixed positive charge and hot-electron immunity for nitrided oxides," *IEEE Electron Device Lett.*, vol. 18, pp. 471–473, 1997.

# SILICON PROCESSING

# FOR

# THE VLSI ERA

## VOLUME 3:

## THE SUBMICRON MOSFET

**STANLEY WOLF Ph.D.**
Professor, Department of Electrical Engineering
California State University, Long Beach
Long Beach, California

**LATTICE PRESS**

Sunset Beach, California

Published by:

Lattice Press.
Post Office Box 340
Sunset Beach, California 90742, U.S.A.

Cover design by Roy Montibon, Visionary Art Resources, Inc., Los Angeles, CA.



**Library of Congress Cataloging in Publication Data**
Wolf, Stanley

Silicon Processing for the VLSI Era
Volume 3 : The Submicron MOSFET

Includes Index
1. Integrated circuits-Very large scale integration. 2. Silicon. I. Title

ISBN 0-961672-5-3

9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

# CONTENTS

8. If the circuit application requires that no more than 10 pA leakage flows between the p
and n-well in a CMOS technology, and we have obtained the process and device data from Fi
6-57a. What design rule should be specified to provide for the minimum $n^+$-to-n-we
separation?

# CHAPTER 7

## THIN GATE OXIDES:

## GROWTH and RELIABILITY

Gate oxide thickness has grown smaller with each generation of MOS ICs. Table 7-1 shows the evolution of MOSFET scaling from 1977 to 1991.[1] As the gate length shrank from 3.0 μm in 1977 to 0.6 μm in 1991, the gate oxide thickness went from 70 nm to 13.5-12.5 nm. During the shrinkage from 3.0 μm to 0.6 μm technologies, the power-supply voltage remained at 5V. At gate lengths below 0.5 μm, however, it is projected that the power supply voltage will have to be reduced. In fact, a reduction to 3.3V in some 0.5 μm CMOS technologies has already been implemented. The decrease in power-supply voltage is needed because the gate oxide thickness in such deep-submicron MOSFETs is 10 nm or less. As will be discussed, the oxide electric fields produced by 5V cause severe reliability problems in oxide films thinner than 10 nm.

The thickness of gate oxides is made as small as possible because the drive current in MOSFETs increases with decreasing gate oxide thickness. Since this trend extends to gate oxide thicknesses below 5 nm, it seems apparent that even thinner oxides will be specified as MOSFETs are scaled below 0.5 μm (see Table 5-3 in Section 5.6.1).

**Table 7-1 Evolution of MOSFET Technology Since 1977[1]**

| Year of Introduction | 1977 | 1979 | 1982 | 1985 | 1988 | 1991 |
|---|---|---|---|---|---|---|
| Gate Length (μm) | 3 | 2 | 1.5 | 1.1 | 0.9 | 0.6 |
| Minimum Feature Size (μm) | 3 | 2 | 1.5 | 1 | 0.7 | 0.5 |
| $V_{CC}$ (V) | 5 | 5 | 5 | 5 | 5 | 5 |
| MOSFET Gate Oxide (nm) | 70 | 40 | 25 | 25 | 20 | 13.5 |
| Junction Depth (μm) | 0.6 | 0.4 | 0.3 | 0.25 | 0.2 | 0.15 |
| NMOS $I_{Dsat}$ @ $V_{GS}$ = 5V (mA/μm) | 0.1 | 0.14 | 0.23 | 0.27 | 0.36 | 0.64 |
| PMOS $I_{Dsat}$ @ $V_{GS}$ = 5V (mA/μm) | | 0.06 | 0.11 | 0.14 | 0.19 | 0.31 |
| Gate Delay @ FO = 1 (ps) | 800 | 350 | 250 | 200 | 160 | 90 |

Furthermore, thin oxides (and shallow junction depths) are needed to control short-chan -nel effects; i.e., thin gate oxides permit the gate to retain strong control of the channel charge, and shallow junctions keep the drain field from extending far into the channel.

Table 5-4 in chap. 5 gives an estimate of MOSFET technology scaling made by Hu in 1993.[1] As the gate length is decreased and approaches 0.1 $\mu$m it is predicted that gate oxides of only 4 nm thickness will be used (albeit at a power supply voltage of 2.2V). An example of a deep submicron CMOS technology (0.25 $\mu$m) is given by Chang et al. of IBM, in which $t_{ox}$ = 7 nm and $V_{DD}$ = 2.5V.[2] Other studies predict that even thinner oxides will be needed in memory technology than in logic CMOS technologies. For example, in 64 MB DRAMs the capacitor oxide thickness is estimated to be below 10 nm, and in EEPROMs tunneling oxides are even thinner than those used in DRAMs.

While the benefits of using thin oxides are apparent, there are two key issues which must be satisfied to permit them to be utilized: (1) it must be possible to grow such thin oxide layers precisely and uniformly across a wafer; and (2) MOSFETs with such thin oxides must exhibit adequate reliability characteristics under normal circuit operating conditions. These two issues are the subject of this chapter.

## 7.1 GATE OXIDE CHARACTERISTICS NEEDED FOR SUBMICRON MOSFETS

The desired characteristics of MOSFET gate oxides are the following:

1. The grown gate oxide thickness must closely match the thickness specification of the MOSFET device design.

2. The specified oxide thickness must also be sufficiently uniform across the entire wafer, and from wafer to wafer, and from run-to-run.

3. The gate oxide film and the $Si/SiO_2$ interface must exhibit adequately small values of charge in the oxide and at the $Si\text{-}SiO_2$ interface (i.e., low $Q_f$, $D_{it}$, $Q_{ot}$ and $Q_m$ values - see chap. 3).

4. The dielectric breakdown strength of the oxide must be sufficiently high (e.g., >8 MV/cm), implying that the film is pinhole free and contains a negligible number of defects that would lead to oxide breakdown at lower electric fields.

5. The oxide film must exhibit a sufficiently long lifetime under normal operating conditions (i.e., $t_{BD}$ is adequate). This characteristic is related to #4.

6. The oxide should exhibit high resistance to hot-carrier damage (i.e., device degradation should be low in the face of CHE or DAHC stressing).

7. If the oxide is to be used in a symmetric CMOS technology (i.e., in which both $p^+$ and $n^+$ poly are used), the oxide film needs to be resistant to the penetration of boron at the process temperatures used after gate doping. (Note that this topic will be discussed in chap. 9.)



From the list of desired gate oxide characteristics given above, we see that thin oxide reliability issues fall into two categories:

a. Dielectric breakdown (which causes catastrophic failure of device).

b. Hot-carrier-injection degradation (as a result of which device characteristics are degraded).

In this chapter we cover gate oxide breakdown, and in chapter 9 we describe device degradation due to hot-carrier injection. However, here we will also cover all the mechanisms by which charge is injected into the gate oxide (section 7.3), including injection by "hot" as well as by "cold" carriers.

## 7.2 PHYSICAL PICTURE of $SiO_2$ and the $Si/SiO_2$ INTERFACE

Before setting out to describe the technology of growing thin gate oxides and their reliability, it is useful to discuss the physical and chemical properties of thermally grown $SiO_2$ and the physical picture of the $Si/SiO_2$ interface. We cover these topics in this section.

### 7.2.1 Physical and Chemical Properties of $SiO_2$

$SiO_2$ (silica) can be found in crystalline, vitreous, or amorphous forms. When the atomic structure of silica exhibits long-range order it is said to be in a *crystalline* form. Many varieties of crystalline silica exist, including quartz, crystobalite, coesite, etc. However, the type of silica encountered in silicon ICs (i.e., thermal silicon dioxide films grown under normal conditions) is not crystalline. Instead, it is silica in its *vitreous* (or *glassy*) form. Vitreous solids do not exhibit long range order. Instead, their structure is highly ordered only over short ranges. This glassy state of $SiO_2$ is also often referred to as *fused silica*. If no short-range order is exhibited by the silica, its structure is termed *amorphous*. Amorphous $SiO_2$ films may be encountered in oxides deposited under poorly controlled conditions. Since we are interested in thermally grown $SiO_2$ most of our discussion will focus on the glassy state of $SiO_2$ (fused silica).

The basic structural unit of silica is centered around the *structural formula* ($SiO_4$). The spatial arrangement of Si and O atoms is due to their respective valence (+4 and -2), their relative sizes, and their electrostatic interactions. These factors give rise to elementary $SiO_2$ cells of tetrahedric configurations. That is, a silicon atom (with a valence of +4) is located at the center of the tetrahedron, with oxygen atoms ($O^{2-}$) at each of the corners (Fig. 7-1a).[3] In crystalline $SiO_2$ (quartz), each oxygen atom belongs to two tetrahedra and is thus bonded to 2 silicon atoms. Such oxygen atoms are then known as *bridging oxygen atoms*. (Fig. 7-1b).[3] In vitreous or amorphous $SiO_2$ some of the vertices of the tetrahedra have *nonbridging oxygen atoms*, meaning they are not shared between two tetrahedra (Fig. 7-1d). The greater the ratio of bridging to non-bridging oxygens, the greater the cohesiveness of the $SiO_2$ structure.

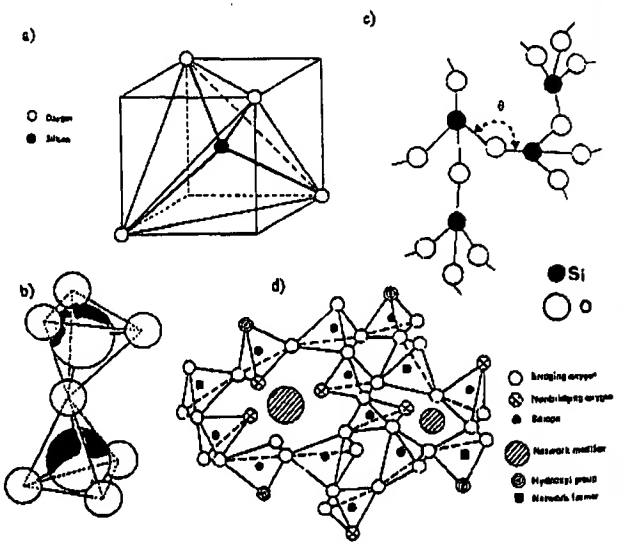


Fig. 7-1 (a) The basic structure of $SiO_2$.[18] (b) Three-dimensional representation of two neighboring $SiO_4$ cells, bridged by an oxygen atom. (c) Two-dimensional lattice representing vitreous $SiO_2$.[18] (d) The structure of thermally grown $SiO_2$ showing bridging and non-bridging oxygen atoms and dopant (i.e., network modifier) atoms.[112] (© IEEE 1965).

The interatomic distances (from the center of one atom to the center of its nearest neighbor) have been measured and found to have the following mean values: 1.62Å for the Si-O bond; 2.27Å for the O-O distance; and 3.12Å for the Si-Si distance. Note this implies that a 10 nm gate oxide is only 40-50 atomic layers thick!

The structure of fused silica is a continuous random network of tetrahedra, where the Si-O-Si bond angle varies from 110° to 180°, with the most probable value being 144°. A two-dimensional representation of the vitreous $SiO_2$ lattice is shown in Fig. 7-1c).[3]

The density of thermal $SiO_2$ (fused silica) is 2.20 g/cm$^3$ which is smaller than that of quartz (2.65 g/cm$^3$). In fused silica only 43% of the lattice space is occupied, making its structure much more open than that of quartz. Consequently, a large variety of impurity atoms can easily enter this oxide network and diffuse through it. Since it is difficult to directly assess the density of thin films, the refractive index of $SiO_2$ is usually measured instead, and the density is then inferred from it. A refractive index of 1.460 corresponds to a density of 2.20 g/cm$^3$. Generally, CVD oxides exhibit smaller densities than thermally grown $SiO_2$, and worse electrical and material properties correspond to less dense films. Hence, measurements of the index of refraction also provide a method for rapidly comparing the characteristics of deposited oxides.



## 7.2.2 The Si/$SiO_2$ Interface

The characteristics of the Si/$SiO_2$ interface, as well as the $SiO_2$ bulk characteristics, play an important role in the functioning of the gate dielectric. Hence, a discussion of the structure of this interface is also useful. In chap. 3 we introduced the concept of surface states and interface trapped charge, and their impact on the threshold voltage of MOSFETs. Here we consider these topics in further detail. We recall that surface states are extra allowed energy states for electrons that are present at the semiconductor surface, but not within the bulk.

The term *surface states* was coined by Tamm in 1932.[4] Using quantum mechanical calculations he showed that new electronic energy states arise in a crystal if the lattice is terminated at a surface. Such new states are confined to the region very close to the surface. He also calculated that these states will arise at interfaces (such as the Si/$SiO_2$ interface), as well as at free surfaces. Each of the states is associated with a single atom at the surface. Hence, an electron occupying one of these states is localized (i.e., forced to remain in a restricted region of space centered on that atom). Since such states thus effectively trap free carriers at the surface, they are also referred to as *interface traps*. The charge per unit area stored in the traps is symbolized by $Q_{it}$.

Although the theoretical basis of interface traps is well accepted and models exist that accurately detail the electrical behavior of the traps, the *physical origin* of these surface states has not been totally worked out. The weight of experimental evidence supports the view that surface states arise primarily from unsatisfied, or *dangling* bonds at the silicon surface (note also that such a silicon atom can be viewed as a *trivalent Si atom*).

From a qualitative perspective, the dangling bond model can be visualized with the aid of Fig. 7-2a. Here it can be seen that if a Si lattice is abruptly terminated along a given plane to form a surface, one of the four bonds of each Si atom at the surface is left dangling. There will thus be extra states on this surface because the energy field of the crystal is one-sided (i.e., the electrons in the surface region are bonded only from the side directed toward the bulk). It is plausible that the thermal formation of $SiO_2$ can tie up these surface Si bonds, and that along a perfect Si/$SiO_2$ interface (Fig. 7-2b) all such bonds could be tied up. In this case, surface states at the interface would be suppressed.[3] However, it is also plausible that such oxidation might not tie up all the bonds (Fig. 7-2c). If even a very small fraction of the number of dangling bonds was left unsatisfied, a significant number of surface states could exist. For example, on a (100) Si surface there are $6.8 \times 10^{14}$ Si atoms per cm$^2$. If oxidation left 1/1000 of these bonds dangling, and each of them gave rise to a surface state, the density of interface trapped charges would be $6.8 \times 10^{11}$ cm$^{-2}$ (assuming one electronic charge is associated with each energy state). If a gate oxide thickness of 20 nm was being used, this would

cause a threshold voltage shift of 0.63V. This indicates that if the dangling bond model correctly describes the origin of interface states, then only a relatively small number of residual dangling bonds can significantly perturb the device characteristics. Crystal defects near the surface or foreign atoms bonded at the surface are other interface perturbations that have been identified as possible sources of surface states.

As proposed by Deal,[6] surface states have the following properties:

1. A surface state can trap or emit a carrier.

2. The energy levels of these states are located within the forbidden gap.

3. The surface state can be of either type: donor or acceptor. The classification is made in a manner analogous to bulk dopant atoms. That is:

   a. A monovalent *donor* trap possesses two charge states +1 and 0. Its charge is positive (+1) when the trap is empty (i.e., its energy level $E_T$ is above $E_F$), but is neutral (0) when full (when $E_T$ is below $E_F$).

   b. A monovalent *acceptor* trap possesses two charge states 0 and -1. When the trap is empty its charge is neutral (0) (i.e., its energy level $E_T$ is above $E_F$), but when the trap is full its charge is negative (-1) (i.e., when $E_T$ is below $E_F$).

4. Because they are mostly due to a disruption of the periodicity of the crystal lattice (and only less frequently caused by the presence of foreign impurities), interface states are not associated with well-defined energy levels. Instead, interface states are distributed throughout the entire band gap at closely spaced levels.

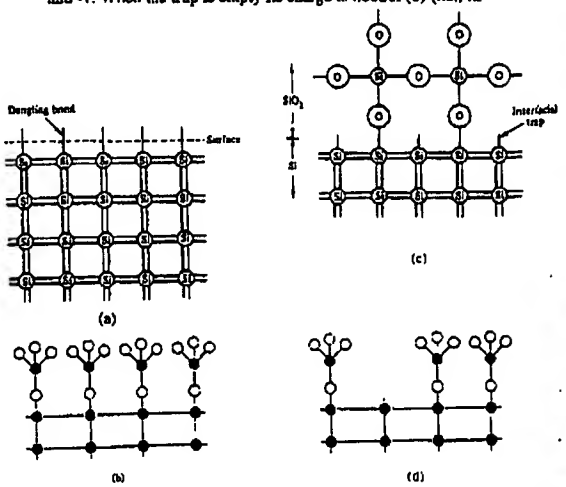


Fig. 7-2 Physical model for the interfacial traps. (a) "Dangling bonds" which occur when the Si lattice is abruptly terminated along a given plane to form a surface. (b) Post oxidation perfect interface. (c & d) Post oxidation dangling bonds that become interfacial traps.[11a] Reprinted with permission of the publisher, the Electrochemical Society.



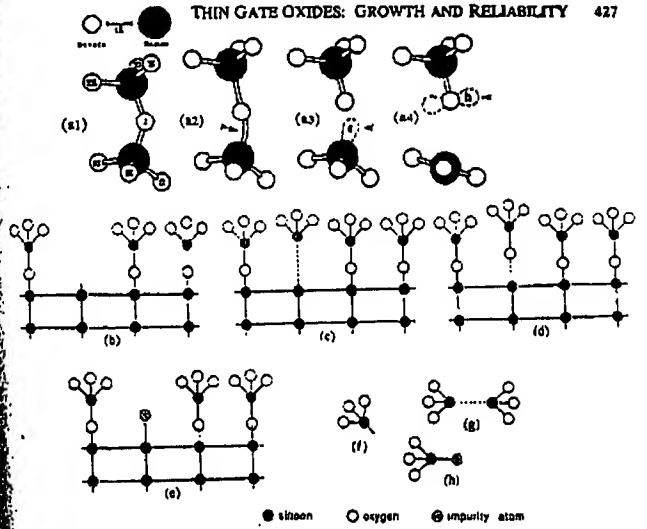


Fig. 7-3 (a) Strained and broken bonds: (a1) normal Si-O bonds; (a2) strained bond in vitreous silica ; (a3) broken bond represented with a trapped electron; (a4) broken bond represented with a trapped hole. Possible interface defects. (b) Si dangling bonds. (c) Si-Si stretched bond (or oxygen vacancy). (d) Si-O stretched bond. (e) impurity at interface. Possible oxide bulk defects: (f) Si dangling bond; (g) Si-Si stretched bond (or oxygen vacancy). (h) impurity in $SiO_2$.[3] (© IEEE 1988).

Figure 7-3 represents the bond stretching and broken bond model of the $Si/SiO_2$ bulk of Verwey[12] and of the $Si/SiO_2$ interface of Sakurai and Sugano.[3] That is, Fig. 7-3a depicts the stretched and broken Si-O bonds in the bulk, and Figs. 7-3b-d at the interface and Figs. 7-3e-h the possible defects in the oxide bulk. In Fig. 7-3b, the dangling Si bond at the surface is depicted as an *interfacial trivalent Si atom*, while in Fig. 7-3f, a dangling Si bond in the bulk is portrayed as a *trivalent Si atom*. Trivalent Si atoms introduce a deep trap level near the Si midgap when located at the interface (see Fig. 7-4) and a deep trap level in the $SiO_2$ bandgap when located in the

oxide bulk (1 in Fig. 7-4). The weak Si-Si* and Si-O stretched bonds (Figs. 7-3c and 7-3d) at the interface introduce a continuum of deep trap levels, or if such stretched bonds exist in the bulk (Fig. 7-3g), they produce shallow trap levels. If a dangling Si bond is tied up with an impurity (most commonly H or OH) this is thought to produce an electron trap, while the oxygen vacancy (or weak Si-Si stretched bond) is considered to be the precursor of the hole trap. Thus, as we shall see, processes that decrease the electron trap densities tend to increase the hole trap densities.

An anneal in a hydrogen ambient (100% $H_2$ or 4%-$H_2$ forming gas) at approximately 450°C is normally the final step prior to assembly and packaging of the ICs. This step is thought to allow hydrogen to penetrate the gate oxide and tie up the remaining dangling bonds at the $Si/SiO_2$ interface not tied up by thermal oxidation (as suggested in Fig. 7-5). However, the Si-H or Si-OH bond can be easily broken by injected hot electrons, giving rise to interface traps. Thus, hydrogen introduced in the post-metal anneal or other process steps (i.e., during the steam reflow of an interlevel BPSG layer, or during the deposition of silicon nitride), can increase hot-electron degradation in a MOSFET. Consequently, it has been suggested that the ideal gate dielectric and all subsequent dielectric layers should contain as little hydrogen as possible to make them hot-electron resistant.

The defects in the oxide bulk and at the interface are believed to be responsible for:

1. Oxide integrity degradation (which may cause catastrophic failure).

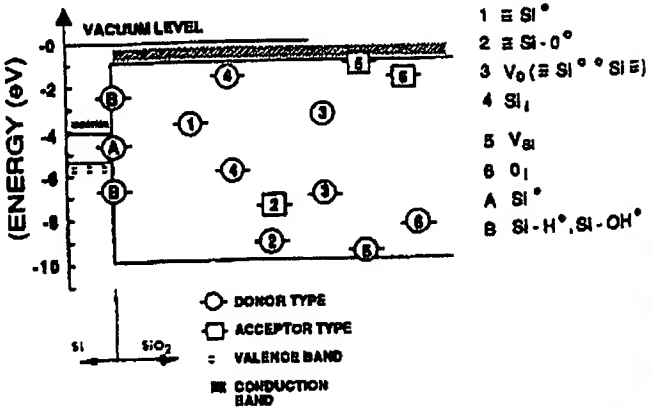


Fig. 7-4 Energy band diagram showing the energy levels of the various defects in the oxide-silicon interface and in the gate oxide.

* The weak Si-Si bond is more frequently referred to as an *oxygen vacancy*.



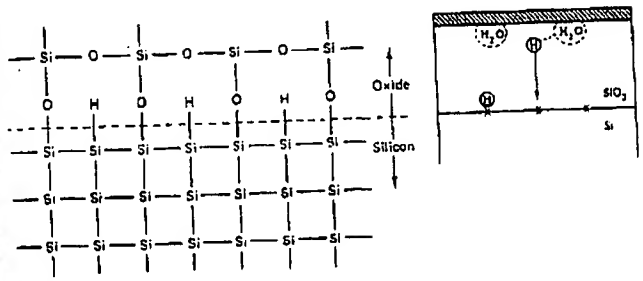


Fig. 7-5 How hydrogen can passivate the dangling bonds at the silicon surface shown in Fig. 7-2c and reduce the interface trap density. From D.L. Pulfrey and N. G. Tarr *Introduction to Microelectronic Devices*, p. 222. Copyright 1989. Prentice-Hall. Reprinted with permission.

2. Hot-carrier induced degradation, which causes some of the device operating characteristics to be degraded (e.g., $V_T$ and $g_m$), but the device may still work.

## 7.3 CARRIER INJECTION IN THE $Si/SiO_2$ SYSTEM

Failures in MOSFETs associated with the gate oxide (both catastrophic failures and device-characteristic-degradation failures) usually involve carriers injected into the gate oxide. A variety of carrier-injection mechanisms have been proposed for the $Si/SiO_2$ system. Here we review the basic physics of the carrier injection phenomena in MOSFETs as useful background material for the discussion of gate oxide breakdown (this chapter), and hot-carrier degradation (chap. 9).

Carriers can be injected either over an energy barrier (hot-carrier injection), or through a barrier (tunneling by cold carriers). The most important injection phenomena are:

1. Channel hot electron (CHE) injection — Hot carrier injection phenomena
2. Drain avalanche hot carriers (DAHC) injection

and

3. Injection by Fowler-Nordheim tunneling — Cold carrier injection phenomena
4. Injection by direct tunneling.

Hot carrier injection implies that the injected carriers are no longer in equilibrium with the lattice at the point of injection (i.e., they possess far more kinetic energy than they would normally acquire from the ambient lattice temperature). They gain such energy

# Improved Hot-Electron Reliability in High-Performance, Multilevel-Metal CMOS Using Deuterated Barrier-Nitride Processing

W. F. Clark, T. G. Ference, S. W. Mittl, J. S. Burnham, and E. D. Adams

*Abstract*— **Deuterated barrier-nitride films and anneals in a deuterium ambient prior to first-metal have been incorporated into a conventional high-performance CMOS process and subjected to subsequent processing through five levels of metal. Device hot-electron stress results confirm that, even though some initial relaxation of the transistor lifetime improvement is observed with further hot processing, significant lifetime improvement can be achieved through full wafer processing through five levels of metal. The barrier-nitride acts as a reservoir for deuterium, maintaining high concentrations in the device regions through further processing. These results support the efficacy of using a deuterium reservoir to achieve a hot-electron-hardened transistor.**

*Index Terms*—**Annealing, deuterium, hot-carriers, MOSFET's, $ND_3$, semiconductor processing, $SiD_4$.**

## I. Introduction

CMOS transistor lifetime degradation from the hot-electron mechanism continues to be a reliability concern as the pace of geometric scaling exceeds the rate of power-supply reduction [1]. This trend, and the additional requirement to interface the lower core voltages to previous generation's higher power-supply circuits, have led to a resurgence in activities to lower the susceptibility of the device to hot-electron degradation. Several years ago, the use of deuterium was proposed to reduce NMOS-transistor susceptibility to hot-carrier degradation by hardening the silicon/silicon-dioxide interface [2]. Recent studies have shown that the replacement of standard hydrogen-based post-metal anneals with deuterium-ambient anneals have yielded significant improvements in hot-electron lifetimes [3]. Subsequent studies have indicated that this anneal effect may not be as readily achieved in device structures incorporating nitride sidewall spacers [4]. Attempts to anneal in deuterium prior to first-metal have shown significant improvement; however, with further thermal processing, the effect is inherently unstable [5]. Higher-temperature and longer-duration anneals, with higher deuterium concentration, at the final-metal level, have been explored with effective results, but are not viable in technologies with barrier-nitride layers [6]. These longer-time anneals also significantly increase process cycle time and their higher temperature can degrade the metal-resistance properties and lead to increased metal voiding [7].

We have previously proposed that deuterium and hydrogen undergo an exchange reaction by which much of the deuterium at the silicon/silicon-dioxide interface becomes diluted from the abundance of hydrogen introduced at other points in the process [8]. One solution would be to introduce a deuterium-rich film above the polysilicon, which functions both as a deuterium reservoir and a barrier to hydrogen entering from subsequent processes. Such an approach, using deuterated ammonia during barrier-nitride layer formation, demonstrated a 6× improvement in hot-electron lifetime at first-metal.

In this letter, we report on an improved barrier layer/reservoir structure formed by replacing all the hydrogen in the barrier-nitride formation with deuterated ammonia ($ND_3$) and deuterated silane ($SiD_4$). All anneals subsequent to the barrier-nitride formation but prior to contact fill are conducted in deuterium. After contact fill, conventional back-end-of-line (BEOL) processing is completed using hydrogen-based anneals; the devices are then processed through five levels of metal to determine the effectiveness of the deuterium reservoir.

***Authors' Note: Ion implantation into films with high deuterium concentration can, under certain conditions, cause a significant release of neutrons. These results will be reported elsewhere. Controls should be in place to prevent implantation of these deuterated films.***

## II. Experimental

MOSFET's processed in a conventional 2.5-V, 0.25-$\mu$m CMOS technology were fabricated for this study. The gate oxide was 5.0 nm. The control samples were processed with conventional nitride films and 5% hydrogen anneals. The experimental samples were processed with a deuterated barrier-nitride layer and annealed in a 10% deuterium ambient prior to nitride formation, at contact-level dielectric densification and at contact liner. The deuterated nitride film was formed by replacing silane and ammonia with $SiD_4$ and $ND_3$ in a plasma-enhanced chemical vapor deposition (PECVD) tool. Process differences are summarized in Table I. A cross section of the device structure used in these experiments is shown in

Mauscript received February 12, 1999; revised June 9, 1999.

The authors are with IBM Microelectronics Division, Essex Junction, VT 05452 USA.

Publisher Item Identifier S 0741-3106(99)07989-6.

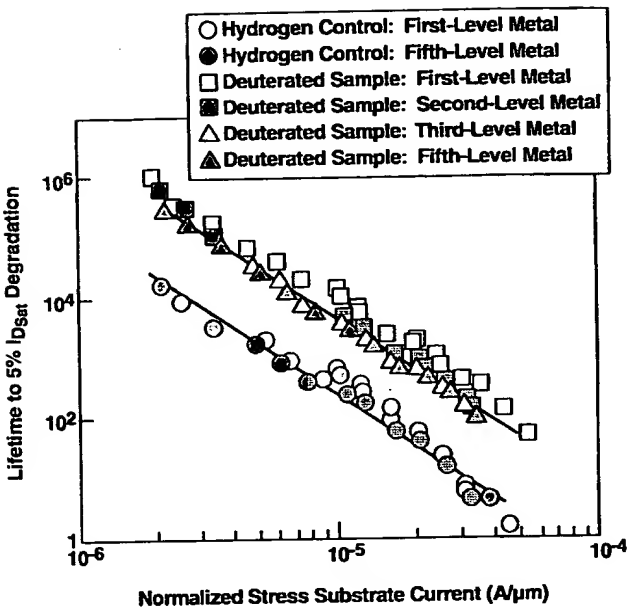


Fig. 1. Lifetime to a 5% degradation in saturated drain current measured in the reverse polarity as a function of normalized stress substrate current for deuterium-processed transistors and hydrogen-processed controls.

TABLE I
SUMMARY OF PROCESS DIFFERENCES

| Process Step | Control | Experimental |
|---|---|---|
| Post Silicide Anneal | 5% $H_2$ / 95% $N_2$ 550°C, 0.75 hr | 10%$D_2$ / 90% $N_2$ 400°C, 1.0 hr |
| Barrier Nitride Deposition | $NH_3$, $SiH_4$ | $ND_3$, $SiD_4$ |
| Contact Level Dielectric Densification Anneal | 5% $H_2$ / 95% $N_2$ 600°C, 0.5 hr | 10% $D_2$ / 90% $N_2$ 600°C, 0.5 hr |
| Contact Liner Anneal | 5% $H_2$ / 95% $N_2$ 550°C, 0.5 hr | 10% $D_2$ / 90% $N_2$ 550°C, 0.5 hr |

the Fig. 2 inset. Both samples were processed through five metallization layers. Conventional hydrogen-based post-metal anneals were conducted on both samples after each metal layer was defined. Hot-electron stressing was conducted after the first, second, third, and fifth (final) metal layers were annealed. The hot-electron stress was conducted at maximum substrate current. Electrical channel lengths as short as 0.12 $\mu$m were stressed. Changes in saturated drain current were monitored as a function of time and substrate current. Film composition was analyzed using hydrogen forward scattering (HFS) and secondary ion mass spectroscopy (SIMS).

## III. RESULTS AND DISCUSSION

Fig. 1 compares the lifetime to a 5% degradation in transistor reverse-saturated drain current as a function of normalized stress substrate current for both the experimental and control FET at several metal levels. A 5% degradation criterion is used because of the limited magnitude of shift on the deuterated transistors. At first-metal, we observed a 28× improvement in transistor lifetime. The lifetime improvement degraded to approximately 18× at second-metal and, by third-metal, stabilized to 13×. We believe this relaxation is caused by

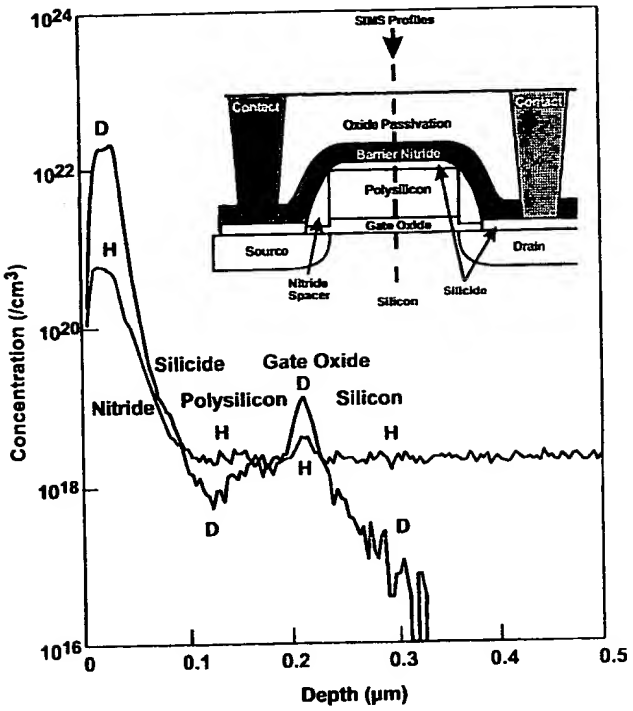


Fig. 2. SIMS profile of the deposited barrier-nitride layer and the transistor gate-oxide regions for the experimental process following contact-level dielectric densification. Inset is a cross section of the device structure used in these experiments.

TABLE II
HYDROGEN FORWARD SCATTERING MEASUREMENTS OF ELEMENTAL CONCENTRATIONS FOR PECVD NITRIDE DEPOSITED WITH $SiH_4$:$NH_3$, AND $SiD_4$:$ND_3$

| Formation Gases | Si | N | H | D | D/(D+H) |
|---|---|---|---|---|---|
| Control $SiH_4$:$NH_3$ | 0.39 | 0.51 | 0.1 | 0 | 0% |
| Experimental $SiD_4$:$ND_3$ | 0.38 | 0.51 | 0.01 | 0.1 | 91% |

diluting the interfacial deuterium with the residual hydrogen in the undeuterated components of the device structure, or by the infusion of hydrogen through the barrier or the contact openings. No significant change in lifetime is observed in the control (hydrogen only) cell from first-metal to last-metal stressing. The pre-nitride anneal in deuterated forming gas can affect final lifetime improvement by 2×.

HFS and SIMS analyses of the nitride film prior to the metallization steps were conducted to determine the relative concentrations of hydrogen and deuterium in the film; the HFS results are given in Table II. The SIMS profiles are shown in Fig. 2 for the deuterated barrier-nitride and the transistor gate-oxide regions after contact-level dielectric densification. The total concentration of hydrogen or deuterium for the control and experimental films, as determined by HFS, was 10 and 11 atomic percent, respectively, while 91% (HFS) to 97% (SIMS) of the hydrogen in the experimental film is replaced by deuterium. The SIMS profile suggests that the effective deuterium concentration in the gate-oxide region after nitride deposition and contact-level dielectric densification was 80%

(by integrated dose). This is consistent with the 28× lifetime improvement observed at first-metal [8].

## IV. Conclusion

Previously, CMOS transistors, incorporating nitride-sidewall spacers and a barrier-nitride layer, were reported to be resistant to improvements by deuterium annealing. We have improved the hot-electron reliability of high-performance CMOS transistors by a combination of a fully-deuterated barrier-nitride and pre-metal anneals in deuterium. This approach is advantageous because it is a highly manufacturable methodology with no additional process steps or cycle time. The improvement also extends to CMOS technologies with five levels of metal with no change to the BEOL processing.

## Acknowledgment

The authors thank J. Panton, M. Triplet, and T. Hetrick for their help in modifying tooling, making it possible to carry out these experiments; D. Mosher and J. Moffatt for depositing the deuterated nitride films; J. Panton and M. Bushey for their help in processing wafers with deuterated gases; A. Sykes for hot-carrier stress assistance; and J. Ziegler, T. Zabel, and M. Zaric for radiation measurements.

## References

[1] H. Momose, S. Nakamura, T. Ohguro, Y. Yoshitomi, E. Morifuji, T. Morimoto, Y. Katuma, and H. Iwai, "A study of hot-carrier degradation in n- and p-MOSFET's with ultrathin gate oxides in the direct tunneling regime," in *IEDM Tech. Dig.*, Dec. 1997, pp. 453–456.

[2] B. S. Lisenker and R. J. Roddy, "Semiconductor device comprising deuterium atoms," Int. Patent Applicat. WO 94/19829, 1994.

[3] J. W. Lyding, K. Hess, and I. Kizilyalli, "Reduction of hot-electron degradation in metal oxide semiconductor transistors by deuterium processing," *Appl. Phys. Lett.*, vol. 68, no. 18, p. 2526, 1996.

[4] S. Aur, T. Grider, V. McNeil, T. Holloway, and R. Eklund, "Effects of advanced processes on hot-carrier reliability," in *Proc. IEEE Int. Reliab. Phys. Symp.*, 1998, pp. 180–183.

[5] W. F. Clark, T. G. Ference, T. B. Hook, K. M. Watson, S. W. Mittl, and J. S. Burnham, "Process stability of deuterium annealed MOSFET's," *Electron Device Lett.*, vol. 20, pp. 48–50, Jan. 1999.

[6] I. Kizilyalli, G. Abeln, Z. Chen, J. Lee, G. Weber, B. Kotzias, S. Chetler, J. Lyding, and K. Hess, "Improvement of hot-carrier reliability with deuterium anneals for manufacturing multilevel metal/dielectric MOS systems," *IEEE Electron Device Lett.*, vol. 19, pp. 444–446, Nov. 1998.

[7] Y. Sugano, S. Minegishi, H. Sumi, and M. Itabashi, "*In situ* observation and formation mechanism of aluminum voiding," in *Proc. IEEE Int. Reliab. Phys. Symp.*, 1988, pp. 34–38.

[8] T. G. Ference, J. S. Burnham, W. F. Clark, T. B. Hook, S. W. Mittl, K. M. Watson, and L. K. Han, "The combined effects of deuterium anneals and deuterated barrier nitride processing on hot-electron degradation in MOSFET's," *IEEE Trans. Electron Devices*, vol. 46, pp. 747–753, Apr. 1999.